A Role for the Spemann Organizer Gene, *Goosecoid*, in Tumor Metastasis

by

Kimberly A. Hartwell

B.S., Molecular, Cellular, and Developmental Biology
Yale College, 1999

Submitted to the Department of Biology
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

February 2007

Signature of Author: _____
Department of Biology

Certified by: _____ 1/22/07
Robert A. Weinberg
Professor for Cancer Research
Thesis Supervisor

Accepted by: _____
Stephen P. Bell
Professor of Biology
Chairperson of the Biology Graduate Committee

A Role for the Spemann Organizer Gene, *Goosecoid*, in Tumor Metastasis

by

Kimberly A. Hartwell

Submitted to the Department of Biology on February 6, 2006
in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Abstract

The process of invasion and metastasis during tumor progression is often reminiscent of cell migration events occurring during embryonic development. I hypothesized that genes controlling cellular changes in the Spemann organizer at gastrulation might be reactivated in tumors. The Goosecoid homeobox transcription factor is a known executer of cell migration from the Spemann organizer. I found that indeed *Goosecoid* is overexpressed in a majority of human breast tumors. Ectopic expression of Goosecoid in human breast cells generated invasion-associated cellular changes, including an epithelial-mesenchymal transition. TGF-β signaling, known to promote metastasis, induced *Goosecoid* expression in human breast cells. Goosecoid induces the expression of E-cadherin repressor SIP1, scaffolding protein IQGAP1, and PDGF signaling components, all which have independently been implicated in tumor metastasis. Moreover, Goosecoid significantly enhanced the ability of breast cancer cells to form pulmonary metastases in mice. These results demonstrate that *Goosecoid* promotes tumor cell malignancy and suggest that other conserved organizer genes may function similarly in human cancer.

Thesis advisor:
Robert A. Weinberg, Ph. D., Professor for Cancer Research, Department of Biology,
Massachusetts Institute of Technology and Whitehead Institute for Biomedical Research

to

my father
Allan D. Hartwell
in loving memory

June 21, 1945 - September 10, 2003
glioblastoma multiforme

Biographical Note

Kimberly A. Hartwell

Whitehead Institute for Biomedical Research
9 Cambridge Center
Cambridge, MA 02142
phone: (617) 258-5158; email: kimberly.hartwell.br.99@aya.yale.edu

EDUCATION

1999-2006 **Massachusetts Institute of Technology** Cambridge, MA
Ph.D. Candidate in Department of Biology. Degree expected February 2007.

1995-1999 **Yale University** New Haven, CT
B.S. in Molecular, Cellular and Developmental Biology with Distinction in Major.
Conducted independent research culminating in senior thesis.

RESEARCH EXPERIENCE

2000-2006 **Whitehead Institute, MIT Department of Biology** Cambridge, MA
Advisor: Dr. Robert A.Weinberg
Research areas:
Role of epithelial plasticity in breast cancer progression.
The *Goosecoid* developmental gene as a promoter and regulator of metastasis.
Aberrant *Goosecoid* expression in clinical tumors.
Work has involved collaborations with colleagues, training and mentoring of students.

1997-1999 **Yale School of Medicine, Department of Pathology** New Haven, CT
Advisor: Dr. Archibald Perkins
Project title: Cooperation of Erb-b2 and p53-172H in Mammary Tumorigenesis.
Work culminated in senior thesis and two presentations.

Summer 1997 **Tufts University, Department of Biology** Medford, MA
Advisor: Dr. Shuk-Mei Ho
Project title: Inducible Apoptotic Genes in Estrogen-Independent Breast Cancer.

1996-1997 **Yale University, Department of Biology** New Haven, CT
Advisor: Dr. Frank Ruddle
Project title: Hox Gene Evolution from Sea Lamprey Gene Mapping.

AWARDS

Predoctoral Traineeship Award, Department of Defense Breast Cancer Research Program, 2002-2005
Howard Hughes Research Fellowship Recipient, Yale University, Summer 1997
Howard Hughes Research Fellowship Recipient, Tufts University, Summer 1996
Science, Technology and Research Scholar, Yale University, 1996-1998
National Scholar, National Science Foundation Scholars Program, 1995

TEACHING EXPERIENCE

Graduate Teaching Assistant, Cell Biology for MIT undergraduates
Ran two course sections per week, wrote exams and problem sets, graded exams (2001)
Graduate Teaching Assistant, Introductory Biology for MIT undergraduates
Ran four course sections per week, gave lectures on course material, wrote problem sets, graded exams (2003)

LEADERSHIP

Science Program Director, Paul Newman's Hole in the Wall Gang Camp for children with cancer (1999)
Redesigned and ran program, resulting in renewed enthusiasm and a complete revamping of facilities.
Program Co-chair and Instructor, Fertile Ground Science Enrichment in New Haven schools (1995-1999)

SELECTED PRESENTATIONS

Era of Hope Department of Defense Breast Cancer Research Program Meeting, Philadelphia, PA 2005
Colrain Meeting on Cancer, Colrain, MA 2004
Conference on Targets in Chemoprevention and Therapy, Austin, TX 2002
Broad Institute Cancer Program Seminar Series, Cambridge, MA 2004
Mammary Gland Research-in-Progress Seminar Series, Yale School of Medicine 1999
Keystone Symposium, The Microenvironment in Cancer Progression, Banff, Canada (poster) 2005

PUBLICATIONS

Hartwell, KA, Muir, B, Reinhardt, F, Carpenter, AE, Sgroi DC, Weinberg, RA. The Spemann Organizer Gene, *Goosecoid*, Promotes Tumor Metastasis. *PNAS*. In press.

Mani AS, Yang J, Brooks M, Schwaninger G, Zhoul A, Miura N, Kutok JL, **Hartwell KA**, Richardson AL, Weinberg RA. Mesenchyme Forkhead 1 (FOXC2) plays a key role in metastasis and is associated with aggressive basal-like breast cancers. In review.

Gupta, PB, Mani, S, Yang, J, **Hartwell, K**, Weinberg, RA. The Evolving Portrait of Cancer Metastasis. *Cold Spring Harb Symp Quant Biol.* 2005;70:291-7.

Lundberg AS, Randell SH, Stewart SA, Elenbaas B, **Hartwell KA**, Brooks MW, Fleming MD, Olsen JC, Miller SW, Weinberg RA, Hahn WC. Immortalization and Transformation of Primary Human Airway Epithelial Cells by Gene Transfer. *Oncogene*. 2002 Jul 4;21(29):4577-86.

Acknowledgements

I would like to take this opportunity to express my sincere gratitude

to Dr. Robert Weinberg, my thesis advisor and mentor. Thank you for a tremendous training experience. I am most grateful for your genuine care, invaluable scientific input, and enthusiasm for a gene called *Goosecoid*. Thank you, as well, for the many opportunities for professional and personal growth, and for allowing me the independence to test my own strength.

to my thesis committee members, Dr. Richard Hynes and Dr. Jacqueline Lees for your scientific insights and guidance, and for your consistent thoughtful mentoring through career exploration, decision-making, and degree completion. Thank you for your generosity and care.

to Dr. Joan Brugge and Dr. Michael Hemann, members of my thesis defense committee, in appreciation for your giving your time and efforts.

to my dear baymate, Ittai Ben-Porath, there for every last step. For all the times you dropped everything to assist or listen, for the daily hilarity that has been sharing a bay with you, for your incredible generosity with your car. I am so lucky to have had such a mentor and friend in you through these years.

to each of my labmates- my teammates, my most intense and fun-loving lab family- your mentoring and friendship have truly been essential. Thank you especially Tan Ince, Jing Yang, Christina Scheel, Sandy McAllister, Mary Brooks, Rick Lee, Sheila Stewart, Charlotte Kuperwasser, S. A. Mani, Akira Orimo, Bill Hahn, Antoine Karnoub for everything from your careful comments on manuscript drafts and scientific strategy, to your generous assistance during my father's illness. Jeanne Winsten and Ferenc Reinhardt- your generosity with your time and care have also been essential. Thank you again to Tan Ince, Jing Yang and S. A. Mani for project cross-fertilizations and helpful discussions.

to Chris Hickey for your patient assistance with the million fine details and for all the thoughtful extras above and beyond, and to Maria Pavao for your assistance with grant application preparation and other essentials.

to my collaborators Dr. Dennis Sgroi, Dr. George Bell, Dr. Sarah Frew, Dr. Lei Xu, for your insights and efforts, for the fun of it all.

to Dr. Anne Carpenter, collaborator and confidant. For your willingness to live in a matchbox, if that's what it would take, for your generosity and support.

to Betsey Walsh and to Dr. Janice Chang, for the steady stream of information and support. Your care is most appreciated.

to Mary Rockas, for your superb technical assistance and unbridled enthusiasm at the bench. You are already a fabulous scientist.

to my mother and father, Janice and Allan Hartwell, for immersing me in the joy of the outdoors and all things biological from the beginning, for consistently vocalizing your confidence in my abilities, for your tremendous love and support. I am grateful.

to Douglas Hartwell (my big brother) and Helen Norris Hartwell. You have been an immense source of strength, light, and inspiration these years. I love you guys!

to my beloved roomies, my sisters, Megan Higginbotham, Anupama Seshan, Jillian Pesin, for sustaining me in laughter and love through so many things. I will always ache for our years on Albion Street together.

and to our biology homies, Nick Bishop, Andy Tolonen, Enrico Montana, and Greg Liszt. What fun (and stress) we've shared these years.

to Kathryn Sievers and Lauren Willig, my dear kindreds, for (independently!) appreciating the humor of gifting chocolate mice to a mouse experimentalist, and the importance of delving in deeply over tea or the telephone. Thank you both for your love and support. You are my Yale treasures.

to Anna-Karin Robertson for your immediate warmth and openness, for your inspiring passion for atherogenesis, for your regular check-ins from New Haven.

to Gerard J. Ostheimer for your love and all the laughter in the final stretch, for the countless library visits, for the luck of writing my thesis in such a happy and hopeful state of mind.

Table of Contents

Chapter 1

Introduction

Metastasis has significant clinical ramifications

Tumor metastasis is the process by which cancer cells disseminate beyond a primary tumor site, seeding secondary tumors in distant organs. The clinical significance of this process is substantial. Metastasis is known to cause over 90% of cancer-related deaths (Sporn 1996). While the eradication of a primary tumor is far from trivial, patients with one localized tumor are more readily treatable, by surgery and/or localized irradiation, than are patients with numerous lesions at various sites within the body. For this reason, an adequate understanding of the process by which cancer cells spread beyond the primary tumor is critical if we are to make gains in patient outcome. Molecular mechanisms found to drive metastasis could prove invaluable if such knowledge were translatable into treatments that restrict a tumor to its original site.

Numerous barriers of great complexity prevent tumor metastasis

To date, our understanding of the mechanisms driving tumor cell dispersal is limited, in part because of the great complexity of this biological process. Carcinoma cells are derived from epithelial cells, and numerous barriers set in place by normal tissue architecture restrict the dispersal of such cells beyond the epithelium of a given organ. If a tumor cell is to metastasize, it must therefore have the ability to bypass these natural barriers. The barriers to metastasis are thought to be complex, as are the various traits that cells must evolve in order to overcome them (Steeg 2006).

Normal epithelial cells have very limited migratory capabilities and are anchored in place both by cell-cell and cell-matrix attachments. Transformed epithelial cells can traverse the restrictive barrier of extracellular matrix proteins known as the basement membrane by losing

these anchors and acquiring cell motility and invasiveness (Gumbiner 1996). The invasion of tumor cells through the basement membrane and into the surrounding stroma is a significant step in tumor progression, paving the way for tumor metastasis. In fact, whether or not a tumor has broken through the basement membrane is one of the most significant prognostic indicators currently in use in the clinic, as this feature distinguishes a carcinoma from a tumor growing *in situ* (Smith et al. 1984).

The detachment of epithelial cells from their cell-cell and cell-matrix linkages normally results in a type of apoptosis called anoikis, a Greek word meaning "homelessness" (Frisch and Francis 1994). Metastasizing cells must acquire genetic or epigenetic lesions rendering them resistant to such a fate(Steeg 2006).

If tumor cells are to reach distant organs, they must enter, or intravasate, into vessels of the vasculature or the lymphatic system, which mediate their transport (Wyckoff et al. 2000). For tumor cells to travel in this way, they must reach and traverse the walls of such vessels. They must also survive within the environment inside, surviving mechanical shearing forces and evading detection by immune cells (Chambers et al. 2002).

Finally, tumor cells that will generate distant metastases must exit the vasculature, which in some cases requires adhesive interactions with the endothelial cells lining the blood vasculature (Nicolson 1988). Exiting into distant organs is thought to occur by active extravasation out between the endothelial cells lining the vasculature. Tumor cells may also effectively exit the vessel by first proliferating within the vessel. If such proliferation generates a tumor mass that exceeds the size of the vessel, the integrity of the vessel becomes compromised (Chambers et al. 2002).

Once a tumor cell reaches the parenchyma of a distant organ, it must possess still other capabilities if it is to generate clinically relevant metastases. Not only does a metastasizing cell have to remain undetected by various immune cells, but it must survive and adapt to an unfamiliar stromal microenvironment. Metastasizing tumor cells must also manage to proliferate within this foreign environment rather than remain dormant, and, ultimately, must generate a vasculature by way of angiogenesis if the resulting metastases are to grow beyond a minimal size (Folkman 1986; Chambers et al. 2002). This later stage of metastasis is termed colonization, the process by which tumor cells seeded in the parenchyma of distant organs generate macroscopic metastases (Fidler 2002).

If a given tumor cell possesses the characteristics necessary to traverse these various barriers to metastasis, clinically-relevant metastases can result.

The genetic basis of metastasis

It is generally understood that tumorigenesis is a multistep process that resembles Darwinian evolution. Successive rounds of mutation, selection and expansion of specific cell variants within a tissue can give rise to transformed cancer cells (Hanahan and Weinberg 2000). Further tumor evolution, or progression, can give rise to cells that are increasingly self-reliant. In time, the continued accumulation of genetic abnormalities can result in tumor cells with metastatic capabilities (Fidler 2003).

Analyses of the gene expression profiles of bulk tumor cell populations have shown that metastatic propensity is reflected in a 'poor-prognosis' signature that is evident relatively early in multistep tumorigenesis (Weigelt et al. 2005). Tumors whose gene expression patterns do not

contain this signature may never progress to metastasis. These differences likely reflect both cell-of-origin effects as well as early mutation events (Gupta et al. 2005b) and suggest that it is the tumor as a whole that predicts metastatic propensity (Ramaswamy et al. 2003). However, metastatic tumors are known to consist of heterogeneous collections of tumor cells, and rare variants within tumors are especially aggressive (Fidler and Kripke 1977; Kang et al. 2003). These variants may result from additional mutations as well as from epigenetic changes induced by the tumor microenvironment, which will be discussed in greater detail. Kang and colleagues confirmed that highly aggressive and weakly aggressive tumor cell variants could be separated from within a 'poor prognosis' parental cell population (Kang et al. 2003). For these reasons, the metastatic propensity of a given tumor is thought to be dictated by the genetic nature of the tumor in its entirety (a.k.a. whether it fits a 'poor' or 'good' prognosis profile) as well as the nature of rare variants harbored within the tumor (Hynes 2003). This indicates that it is not merely the acquisition of late-stage mutations that causes metastasis but the nature of early events as well (Bernards and Weinberg 2002).

Transformed cells are thought to acquire a metastatic phenotype by evolving the capacity to bypass the natural barriers limiting tumor cell dispersal, as described. The selection of rare variant tumor cells with these capabilities likely occurs both within the primary tumor mass and at distant sites. Selective pressure is provided, at least in part, by these biological barriers (Fidler 2003). Experimental evidence has indicated that colonization, or the creation of macrometastases within a foreign organ, is the main bottleneck in metastasis. Chambers and colleagues have observed that, while the vast majority (83%) of melanoma cells injected into the mouse circulation arrest in the liver, the first capillary bed, only 0.02% of the original cell population could progress beyond this stage to form macrometastases (Weiss 1990). Other studies also

suggest that additional rounds of mutation and cell selection may occur beyond a primary tumor, at the distant metastatic site. For example, it has been demonstrated that tumor cells harbored within the parenchyma of a distant organ can awaken from dormancy upon acquiring angiogenic capabilities, allowing micrometastases to become vascularized (Holmgren et al. 1995). Another study revealed that the genetic abnormalities in single breast tumor cells in the bone differ significantly from that of cells within the primary tumor. This indicates that such tumor cells most likely dispersed quite early during tumor progression and that their genomes continued to evolve at the distant site (Schmidt-Kittler et al. 2003).

Experimental approaches for identifying genes underlying metastasis

As has been previously mentioned, the genetic basis of the metastatic phenotype is thought to be significantly complex, given the numerous and substantial biological barriers that prevent metastasis. Certain molecular signals that help tumor cells overcome these various barriers have only recently been elucidated, and our understanding of the molecular mechanisms behind metastasis remains limited. Nevertheless, to date, several experimental approaches have identified numerous genes underlying the metastatic phenotype.

One especially powerful approach harnesses *in vivo* selection pressures, enabling the isolation of more aggressive subsets of tumor cells contained within a heterogeneous tumor cell population as well as increasingly aggressive derivatives thereof. Subsequent gene expression analyses can identify metastasis genes that may be functional drivers of the more aggressive phenotype. Examples include the use of the experimental metastasis assay as a means of screening for metastasis genes, notably employed by Clark and colleagues, in which especially aggressive melanoma cell variants within a mildly metastatic cell line were isolated following

the injection of the parental cell line into the tail vein of laboratory mice (Clark et al. 2000). Cells injected in this way travel to the first capillary bed they come upon, in this case that of the lung, yet only those cells that can successfully colonize the lung will generate lung macrometastases. Cells that are re-isolated from the metastases that eventually arise in the lungs of such mice can either be re-injected for the selection increasingly aggressive variants or can be examined directly, by mRNA microarray analysis.

Comparisons made in this way between the gene expression patterns in the parental versus the re-isolated tumor cell populations served to identify genes that correlate with the metastatic phenotype (Kang et al. 2003). Such genes can then be subsequently functionally validated as bona fide necessary and/or sufficient mediators of metastasis, as was successfully demonstrated in the Clark study. Variations on this approach have also led to the successful identification of genes that determine tumor cell-target organ compatibility, an important area of study within metastasis research.

Another observation made by Yang et al in our laboratory during the course of the research described herein made use of a set of four mammary cancer cell lines derived from a single spontaneously-arising mouse mammary tumor (Aslakson and Miller 1992; Yang et al. 2004). While the four cell lines form primary tumors that grow with the same kinetics when injected orthotopically into mice, they do not possess equivalent metastatic capabilities. One line generates macroscopic lung metastases, another falls a step short by reaching the lung by way of the vasculature but remains dormant within the lung parenchyma, another cannot extravasate out of the vascular circulation, and a fourth cannot enter the vasculature but remains localized at the primary tumor site. A comparison of the mRNA in primary tumors generated with these lines, by

microarray analysis, identified genes that were subsequently confirmed to be essential for metastasis (Yang et al. 2004).

Other *in vivo* approaches have led to the identification of genes that are essential to tumor invasion and metastasis. Transgenic mouse models of tumorigenesis express activated oncogenes or contain silenced tumor suppressor genes. Such mice are powerful tools, as they often reproducibly exhibit multistep tumorigenesis and can be interbred to examine the cooperativity of various oncogenes (Van Dyke and Jacks 2002). Two examples are the Rip1Tag2 mouse model of pancreatic β-cell carcinogenesis (Perl et al. 1998) and the polyoma middle T oncoprotein (PyMT) model of breast carcinogenesis (Lin et al. 2003).

In vitro assays have also been developed that allow for the modeling of various aspects of tumorigenesis and progression in an especially tractable way. Such assays generally model particular subsets of tumor cell functional phenotypes that are known correlates of metastatic ability *in vivo*. Examples include 3D epithelial cultures, which can be used to examine mechanisms regulating cell polarity, apoptosis, proliferation, and invasion, among other properties (Debnath and Brugge 2005). Other examples include (1) wound closure assays and transwell migration assays, which measure cell migration; (2) transwell invasion assays, which model the migration of cells through a basement membrane (Eccles et al. 2005); and (3) soft agar assays, which can serve as a readout for anchorage-independent growth (Wang 2004).

The analysis of clinical tumor samples can also identify genes underlying metastasis

A key reagent for identifying genes that are important in tumor progression and metastasis is clinical tumor tissue. Such tissue can be obtained with permission from cancer patients, and normal tissue from the relevant type of epithelium can sometimes be obtained for

use as a control. Samples of metastases are only rarely obtained, as the surgical removal of metastases is relatively uncommon. mRNA expression analysis using microarrays and other technologies (e.g. SAGE, differential display) provides a comparison of the gene expression patterns in the tissue types examined (Weeraratna 2005).

One benefit of directly examining clinical tumors is that genes identified as differentially expressed between aggressive and non-aggressive samples, for example, might actually play a functional role in clinical tumors, given that they are identified in a patient-relevant context. However, genes whose expression patterns vary with a clinical phenotype are merely showing a correlative relationship and functional assays are therefore still required to confirm whether or not such genes have functional significance.

Such an examination of clinical tumors was used to identify a 'poor-prognosis' signature present in a subset of tumors, as discussed earlier (Weigelt et al. 2005). However, such examinations of the global gene expression of bulk tumors have met with some controversy. Certain signatures are not always reproducible across different sets of tumor samples and in different labs (Li et al. 2005). This may be in large part because clinical samples are usually a mixture of various cell types, and therefore most analyses are performed using samples that contain tumor cells as well as stromal cells (immune cells, fibroblasts, endothelial cells, etc) in unknown and possibly inconsistent proportions. This complication can limit the power of bulk tissue comparisons, and has prompted the use of laser capture microdissection. Microdissection allows for an accurate comparison of tumor cell populations of known identity, because only the desired subset of cells within a tissue are examined (Ladanyi et al. 2006).

A role for the stromal microenvironment in tumor progression

Importantly, the phenotype of tumor cells is not dictated solely by the genetics and epigenetics of the tumor cells themselves, but also by the effects of the surrounding microenvironment (Fidler 2003). Carcinoma cells grow in close juxtaposition to their stromal cell neighbors, including fibroblasts, immune cells, and cells of the vasculature which they are able to recruit and subvert by mechanisms that are incompletely understood. Stromal cells recruited and altered in this way can generate what is termed a desmoplastic response, which is marked by changes in the composition of the extracellular matrix (Dvorak 1986). Such a recruitment of stromal cells can generate an altered tumor microenvironment that favors tumorigenesis and progression (Elenbaas and Weinberg 2001). Experiments have in fact shown that a desmoplastic response actually promotes tumorigenesis, predisposing adjacent epithelial cells to the eventual acquisition of genetic mutations (Sternlicht et al. 1999).

Most stromal cells present within a tumor are myofibroblasts, or 'activated' fibroblasts, which are a type of fibroblast normally found in wounds during healing (Ronnov-Jessen et al. 1996). Myofibroblasts express large quantities of growth factors, proliferate more rapidly than normal, and are thought to give rise to the tumor-promoting desmoplastic response. Indeed, a large body of research connects the wound-healing response to tumorigenesis and progression, and, consequently, tumors are frequently referred to as wounds that do not heal (Dvorak 1986).

Interestingly, stromal cells are thought to co-evolve along with their tumor cell neighbors, as evidenced by distinct epigenetic differences in normal versus tumor-associated stromal cells that vary with tumor stage (Hu et al. 2005). In a study in our lab by Orimo and colleagues, fibroblasts within a carcinoma were found to acquire expression of SDF-1, which consequently recruit endothelial progenitor cells important for angiogenesis into the primary tumor mass.

SDF-1 can also bind the CXCR4 receptor, which is often abnormally upregulated on the tumor cells themselves, directly stimulating tumor cell proliferation (Orimo et al. 2005). The persistence of these contextual signals can generate increasingly aggressive tumor cell populations over time. However, it is important to note that cancer cells within a tumor are differentially exposed to such environmental influences.

Tumor cells within a tumor mass show variable gene expression, not only according to their genetic and epigenetic differences, but also according to their location within the tumor mass. Distinct differences in the levels and/or subcellular localization of various proteins have been observed in tumor cells at the invasion front of certain tumors, where tumor meets stroma, compared to cells at the center of the tumor. Examples of proteins observed to be upregulated at the invasive edge of tumors include tenascin C, various matrix metalloproteinases, and nuclear β-catenin (Kitadai et al. 1996; Kuniyasu et al. 2000; Brabletz et al. 2001; Beiter et al. 2005). Other proteins are downregulated at the tumor edge, such as E-cadherin (Brabletz et al. 2001). Strikingly, the upregulation of tenascin C, MMPs, and nuclear β-catenin are known correlates of poor prognosis (Jahkola et al. 1998; Ondruschka et al. 2002; Baldus et al. 2004), as is the downregulation of E-cadherin (Berx and Van Roy 2001).

Together, these observations suggest that the cells at the invasion front within tumors may be especially aggressive because of exposure to stromal signals, and that it is this subpopulation of tumor cells that ultimately seed metastases in distant organs (Brabletz et al. 2001). Moreover, the gene expression patterns distinct to this subpopulation of cells may be unstable because these patterns presumably depend on the continued presence of the contextual signals that induced them.

The variation in gene expression between tumor subregions warrants careful study, as it is likely to reveal key mechanisms driving tumor progression. Unfortunately, standard gene expression analyses of clinical or experimental tumors can easily miss these variations. Most often the tissue examined consists of tumor cells from both the inner and outer subregions of a tumor, and subtle gene expression changes localized to the edge of a tumor become obscured when the mRNA of cells throughout a tumor is pooled. This is especially likely to be the case for classes of genes such as transcription factors that can generate functional changes in cellular phenotype with only a relatively modest degree of upregulation. Indeed, a comparison of the gene expression profiles of tumor cells at the edge versus the center of tumors by careful microdissection will likely identify invasion and metastasis regulators.

Evidence of an epithelial-mesenchymal transition (EMT) in cancer

When this thesis work began, a role for stromal cells in tumor progression was becoming increasingly apparent. In thinking about the possible role of the stroma in promoting tumor cell aggressiveness, we were especially intrigued by reports in the literature of cell transdifferentiation occurring within clinical tumors. More specifically, evidence of a transdifferentiation process known as an epithelial-mesenchymal transition (EMT) had been reported (Birchmeier et al. 1996).

By definition, carcinomas derive from the epithelial cells that comprise a given tissue. During an EMT, such cells lose their polarity, change morphologically as a result of cytoskeletal rearrangements, and undergo significant changes in cell-cell and cell-matrix adhesion. These effects are mediated, at least in part, by the downregulation of epithelial cellular components (e.g., E-cadherin and cytokeratins) and by the upregulation of mesenchymal proteins (e.g., N-

cadherin and vimentin) in their stead (Thiery and Sleeman 2006). Epithelial cells that have undergone an EMT in this way acquire properties that are more characteristic of mesenchymal cells (e.g., fibroblasts), including motility and invasiveness. A schematic of the EMT shift is shown in Figure 1.

Moreover, histological examinations of various human primary tumors and metastases had revealed that an EMT is generally restricted to the outer region of an expanding tumor mass (Birchmeier et al. 1996; Brabletz et al. 2001). This restriction of an EMT to the invasion front suggested that the signal(s) that elicit an EMT may be contextual, possibly originating in the stromal cells at the invasion front, where the tumor and stromal cell populations are in close juxtaposition. Moreover, it supported the idea that the EMT transition might be critical for tumor invasion (Birchmeier et al. 1996).

In vitro assays confirmed that the acquisition of these EMT-associated mesenchymal qualities enhances the invasiveness and motility of cancer cells (Birchmeier and Birchmeier 1995). Moreover, at the time we began this work, an EMT was believed to play a significant role in human cancers, including breast, gastric, and colon. Evidence of an EMT in clinical tumors was known to correlate with a poor prognosis (Birchmeier et al. 1996).

During the course of this thesis work, several experimental studies provided additional support for a functional link between EMT and tumor metastasis. For example, certain proteins that can induce an EMT in mammalian mammary epithelial cells, such as Twist and TGF-β, were found to be necessary for the metastatic behavior of tumor cells *in vivo* (Grunert et al. 2003; Yang et al. 2004). FSP1/S100A4, a fibroblast-specific protein, was found necessary for the induction of an EMT by certain cytokines, and the selective depletion of carcinoma cells inducing this protein within a primary tumor resulted in the suppression of metastasis *in vivo*

(Okada et al. 1997; Xue et al. 2003). And, in clinical breast tumors alone, partial or complete loss of E-cadherin has been confirmed to correlate with tumor grade, invasiveness, de-differentiation, and poor prognosis (Berx and Van Roy 2001).

The EMT is essential for numerous normal biological processes

Importantly, the EMT process is not strictly a pathological process. On the contrary, it is known to be essential for normal development and tissue homeostasis. More specifically, an EMT is required for various tissue remodeling events that occur during embryonic development (Thiery 2003), wound healing (de Iongh et al. 2005), and mammary branching morphogenesis (Fata et al. 2004). The concept of an EMT in fact originated from studies of organismal development. The various cell movements that drive gastrulation and other developmental processes, such as neural crest formation and organogenesis, all require an epithelial-to-mesenchymal shift in cellular phenotype (Shook and Keller 2003).

One developmental process that displays especially striking similarities to tumor invasion and metastasis is gastrulation. This critical process generates the three germ layers- ectoderm, mesoderm and endoderm- and establishes the basic organismal body plan by way of highly coordinated cell movements. Gastrulation is initiated by a conserved group of cells originally characterized in *Xenopus laevis* as the Spemann organizer (Niehrs 2004). The organizer is conserved across vertebrates and corresponds to the dorsal blastopore lip in *Xenopus*, the embryonic shield in zebrafish, Hensen's node in birds, and the anterior primitive streak (or node) in mouse (De Robertis 1995).

In higher vertebrates, the organizer cells undergo an EMT at the onset of gastrulation. These epithelial cells break their cell-cell junctions, acquire a mesenchymal morphology, and

ingress into the interior of the embryo, migrating as individual mesenchymal cells (Gilbert 1997). These changes mirror the biochemical and functional changes that define an EMT during tumor progression. Thus it may be that the tumor cell plasticity thought to drive metastasis results from the corruption of an endogenous cell plasticity program essential for normal development that is hardwired into all cells.

Extracellular signals implicated in EMT induction

To date, numerous signaling pathways have been implicated in EMT regulation. Signals observed to be capable of inducing an EMT in epithelial cells *in vitro* include SF/HGF (scatter factor/hepatocyte growth factor), FGFs (fibroblast growth factors), EGFs (epidermal growth factors), TNFα (tumor necrosis factor α), PDGFs (platelet-derived growth factors), members of the TGF-β (transforming growth factor β) superfamily, Wnt/β-catenin signals, and IGFs (insulin-like growth factors) (Figure 1) (Jouanneau et al. 1991; Hoschuetzky et al. 1994; Gilles et al. 1997; Portella et al. 1998). Just as evidence of an EMT in tumors is associated with a poor prognosis, so are many of these EMT regulators (Lin et al. 2000).

While the exact mechanisms by which these signaling pathways induce an EMT remain under intense investigation, many important concepts have emerged. The degree to which various types of cells in various settings undergo a transition from the epithelial to mesenchymal phenotype in response to given signals differs significantly. This indicates that numerous states of partial EMT exist and that the EMT is not simply a binary process (Thiery and Sleeman 2006). Also, there is extensive cross-talk between different EMT-inducing signaling pathways, and thus the orchestration of this switch between the epithelial and mesenchymal states is complex.

Provocatively, at the time we began this thesis work, evidence of elevated TGF-β and Wnt-β-catenin signaling had been detected at the invasion front of certain clinical tumors (Dalal et al. 1993; Brabletz et al. 1998). The upregulation of these signals specifically at the tumor edge supported a role for these signals in inducing the EMT and aggressive cell phenotypes associated with this tumor subregion. Moreover, this observation suggested that these signals were elevated as a result of contextual signals, as previously mentioned.

More specifically, TGF-β1 levels were found to be elevated at the edge of both infiltrating ductal-type breast carcinoma and associated lymph node metastases (Dalal et al. 1993). In a second study, the nuclear localization of β-catenin, a protein in the Wnt pathway, correlated with evidence of an EMT at the invasive edge of both primary colon tumors and metastases (Brabletz et al. 1998). The β-catenin in cells at the center of these tumors was instead observed to be localized to the cell membrane, where it is anchored to E-cadherin as a component of adherens junctions. The translocation of β-catenin to the nucleus was known to be contingent upon both its release from these junctions as well as its protection from APC (adenomatous polyposis coli)-mediated degradation. Nuclear β-catenin binds Tcf/Lef (T-cell factor/lymphoid enhancer factor) transcription factors to activate numerous target genes (Kikuchi 2000). These target genes include known oncogenes *c-myc* and *cyclin D1* (Tetsu and McCormick 1999), as well as EMT-associated genes including *slug*, *fibronectin, vimentin*, and several matrix metalloproteinases (Brabletz et al. 1999; Gradl et al. 1999; Gilles et al. 2003; Sakai et al. 2005). Importantly, the nuclear localization of β-catenin in cancer correlates with a poor prognosis (Lin et al. 2000).

The Spemann organizer gene, *Goosecoid*, as a potential cancer gene

The parallels between gastrula organizer biology and tumor malignancy suggested that common signals may drive gastrulation and metastasis. We therefore directed our attention to the *Goosecoid* (*Gsc*) gene, which encodes a well conserved transcription factor that recapitulates many of the properties of the organizer when ectopically expressed in the amphibian embryo (Cho et al. 1991; Blum et al. 1992; DeRobertis 2004). Intriguingly, elements of the TGF-β superfamily and Wnt/β-catenin signaling pathways, which are known to be upregulated at the invasion front of certain tumors as just discussed, can synergistically induce *Gsc* expression in embryonic cells and are required for Spemann organizer formation (Watabe et al. 1995; Moon and Kimelman 1998; Thiery 2002). These parallels became the basis for the thesis research described herein and they are summarized graphically in Figure 2. For these reasons, we sought to ascertain whether the organizer gene *Gsc* plays a role in neoplastic disease. *Goosecoid* and its encoded protein had not been previously studied in the context of human cancer pathogenesis.

Goosecoid in embryogenesis

Over two decades ago, developmental biologists performed a screen specific for homeobox genes expressed in the *Xenopus* organizer, in an attempt to identify master regulators of Spemann organizer function. *Gsc* was first identified as the most highly expressed homeobox gene in the *Xenopus* organizer by way of this screen (Blumberg et al. 1991). *Gsc* was named as such because the homeodomain of its encoded protein was observed to be similar to that of the Drosophila proteins Bicoid and Gooseberry (Blumberg et al. 1991). Gsc is a member of the paired homeodomain protein family and contains a lysine at position 50 in the homeobox (instead of glutamine as in Hox genes) (Blumberg et al. 1991). Since the discovery of *Gsc*, a closely related gene was discovered and named *Goosecoid-like (Galili et al. 1997).*

Gsc is expressed specifically in the dorsal lip of the *Xenopus* blastopore at the onset of gastrulation. Indeed the expression of *Gsc* serves as a marker of organizer tissue (Cho et al. 1991). All vertebrates examined have been found to express *Gsc* in their organizer region, and invertebrates express *Gsc* at gastrulation as well (Broun et al. 1999; Arendt et al. 2001). Later, days after the initiation of gastrulation, *Gsc* is expressed in the developing head, limbs, and ventral body wall (Rivera-Perez et al. 1995; Yamada et al. 1995). Notably, *Gsc*-like shows a different expression pattern as it is expressed in the developing brain and primordial germ cells (Galili et al. 1998).

The injection of *Gsc* mRNA opposite the *Xenopus* organizer region generates a secondary embryonic axis, demonstrating that *Gsc* can execute some of the organizer properties associated with the dorsal lip (Cho et al. 1991; Yao and Kessler 2001). Importantly, *Gsc* has separately been shown to activate the migratory properties of cells expressing it (Niehrs et al. 1993). In addition, uninjected neighboring cells are recruited to this twinned dorsal axis by the action of secreted Gsc targets, such as Chordin and Frzb-1, a Wnt antagonist (Sasai et al. 1994; Leyns et al. 1997).

A *Gsc* knockout mouse has been generated and mutants die shortly after birth and show various abnormalities but do not have a gastrulation phenotype. This observation came as a surprise and directly contradicted observations in *Xenopus* showing that antagonism of *Gsc* function results in inhibition of both gastrulation movements and dorsal differentiation, generating severe axial defects (Steinbeisser et al. 1995; Ferreiro et al. 1998; Yao and Kessler 2001). The absence of a gastrulation defect in the *Gsc* knockout mouse may result from functional compensation by other genes (DeRobertis 2004). The normal expression patterns and functions of such genes may or may not match those of *Gsc* in murine embryos. *Gsc-like*, a *Gsc* homolog, may partly compensate for *Gsc* (Funke et al. 1997), as may HNF3β/Foxa2, a nuclear

transcription factor shown to be capable of compensating for *Gsc* to some degree (Filosa et al. 1997).

Gsc is able to function both as a transcriptional repressor and activator by way of its binding partners. The organizer-specific secreted protein, chordin, for example, is upregulated by Gsc in *Xenopus* embryos (Sasai et al. 1994). Gsc directly represses other genes, such as Xwnt8, an organizer function antagonist, and Gsc itself in an autoregulatory negative feedback loop (Danilov et al. 1998; Yao and Kessler 2001). Gsc represses target genes by the necessary binding of Groucho (Gro) co-repressor proteins (Jimenez et al. 1999), and target genes are thought to be silenced by both histone deacetylase-dependent and -independent mechanisms (Chen and Courey 2000). Gsc binds target sequences as well as its various binding partners through the near-palindromic DNA sequence TAATCCGATTA (Wilson et al. 1993).

The *Gsc* promoter contains two growth-factor-responsive elements as previously mentioned. A proximal element is activated by Wnt/β-catenin signals and a distal element by TGF-β/Activin/Vg1/Nodal signals (Watabe et al. 1995). More specifically, the proximal element is directly bound by homeodomain proteins Siamois and Xtwn, whose expression is induced by maternal β-catenin (Laurent et al. 1997). The distal element is directly bound by a complex consisting of Mixer and Smad2/4, transcription factors that are activated by TGF-β/Activin/Vg1/Nodal signals (Figure 2) (Germain et al. 2000). Both the proximal and distal elements are necessary for adequate levels of *Gsc* expression in the Spemann organizer and are structurally and functionally conserved between *Xenopus* and mouse (Watabe et al. 1995).

TGF-β signals and Wnt/β-catenin signals are active in distinct regions of the *Xenopus* embryo (throughout the vegetal hemisphere and the dorsal side, respectively) and the site of overlap corresponds to where the organizer forms and Gsc is expressed (Watabe et al. 1995;

Larabell et al. 1997). Siamois, TGF-β signaling, and β-catenin are required for formation of the organizer and embryonic axes and for the expression of organizer genes (Hemmati-Brivanlou and Melton 1992; Heasman et al. 1994; Fan and Sokol 1997). Together, these observations suggest a model in which the spatial overlay of TGF-β signals and Wnt/β-catenin signals specifies the organizer by activating organizer genes such as *Gsc*, which in turn activate additional organizer genes such as chordin (Sasai et al. 1994; Watabe et al. 1995).

In the mouse, the transforming growth factor β (TGFβ), Wnt, and FGF signaling pathways are required for cell migration from the primitive streak (a.k.a. mouse organizer) and for *Gsc* expression at gastrulation (Liu et al. 1999; Sun et al. 1999; Brennan et al. 2001). Activin can induce *Gsc* in early mouse embryos (Blum et al. 1992), as it can in *Xenopus* animal caps in the absence of protein synthesis (Cho et al. 1991). TGF-β and FGF-5 are expressed in the mouse gastrula at the time of *Gsc* induction, yet the localization of the expression of these factors suggests that they may not induce Gsc (Haub and Goldfarb 1991; Hebert et al. 1991; Slager et al. 1991).

Other developmental transcription factors since linked to the EMT and tumor invasion

Since I first began this line of inquiry, other developmental transcription factors, including SNAI1 (Snail), SNAI2 (Slug), SIP1(ZEB2), Twist, δEF1 (ZEB1), and E47 (E12), have been linked to E-cadherin downregulation and tumor cell invasion (Savagner et al. 1997; Batlle et al. 2000; Cano et al. 2000; Comijn et al. 2001; Perez-Moreno et al. 2001; Yang et al. 2004; Eger et al. 2005; Gupta et al. 2005a). E47 and Twist are basic helix-loop-helix proteins. Snail and Slug are zinc finger proteins, as are SIP1 and δEF1 but with the addition of a homeodomain

(Thiery 2002). All have been found to transcriptionally repress E-cadherin by directly binding the E-box elements within the proximal promoter (Hajra et al. 2002) (Eger et al. 2005) (Perez-Moreno et al. 2001; Yang et al. 2004).

During embryogenesis, SNAI1 (Snail), SNAI2 (Slug) and Twist are required for mesoderm formation in *Drosophila* and neural crest development in vertebrates (Thiery and Sleeman 2006), two processes in which an EMT and cell migration are critical (Shook and Keller 2003). E47 expression is specific to the mesoderm in the mouse, suggesting a role for this transcription factor in either the initiation or the maintenance of the mesenchymal state, or both. SIP1 was discovered by way of its interactions with Smads, and is expressed in human tissues that are E-cadherin negative, including in the central nervous system, various muscle types, and in hematopoietic cells (Postigo and Dean 2000).

Other classical developmental pathways, such as the Notch/Jagged and Hedgehog pathways, have also been recently tied to EMT induction. Notch was found to be expressed during development where an EMT occurs, including during cardiac development where its function is in fact essential for endocardial EMT (Timmerman et al. 2004). Moreover, this signal was also found to be necessary for a TGF-β induced EMT in epithelial cells. Integrated Notch and TGF-β signaling downregulates E-cadherin expression by inducing Snail (Timmerman et al. 2004; Zavadil et al. 2004).

The independent discovery of a role for these various developmental genes in the EMT and tumor cell invasion, together with our studies of Gsc, strongly argue that the reactivation of developmental gene programs in human tumors is a potent route to metastasis.

Perspectives

The described similarities between the biology of tumor metastasis and gastrulation suggested that genes controlling cell migration in the embryonic organizer might likewise promote tumor invasion and metastasis. Provocatively, the well conserved transcription factor Goosecoid, which is essential for normal cell migration from the Spemann organizer, was known to be synergistically regulated at gastrulation by signaling pathways that were also known for their prominent roles in tumor invasion and metastasis. I therefore hypothesized that the *Goosecoid* homeobox gene may play a significant role in tumor metastasis, a context in which it had not been previously examined.

In addressing this hypothesis, I sought evidence of *Goosecoid* expression in human tumors in order to ascertain its potential clinical relevance. I used multiple approaches to identify signals that might induce Goosecoid during tumor progression, as well as signals that Goosecoid might function through. I also sought to discern the functional consequences of Goosecoid expression in adult epithelial cells. The results described here strongly support the notion that this embryonic transcription factor can indeed be appropriated opportunistically by human cancer cells, allowing such cells to acquire certain characteristics needed to overcome key barriers to tumor metastasis.

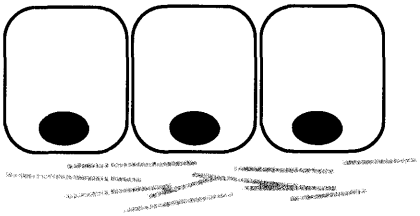
Figure 1. The epithelial-mesenchymal transition (EMT)

Cellular changes associated with an EMT are depicted in the schematic. Examples of epithelial markers that typically are lost as part of an EMT, such as various adherens and tight junction components, are listed, as are various mesenchymal markers that are typically induced. Regulatory pathways that are known to play a role in regulating the EMT are also listed.

EMT-inducing pathways

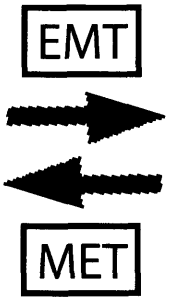
TGF-ß
Wnt/ß-catenin
PDGF
FGF
Notch/Jagged
Hedgehog
TNFα
EGF
IGF
HGF

EPITHELIAL

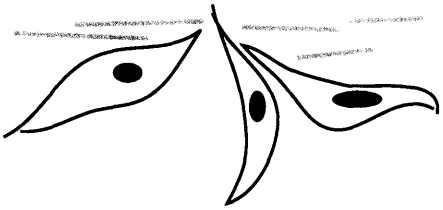


polarized
cobblestone morphology

markers:
E-cadherin
cytokeratins
ZO-1
laminin
claudins
desmoplakin
mucin-1



MESENCHYMAL



nonpolarized
fibroblastic morphology

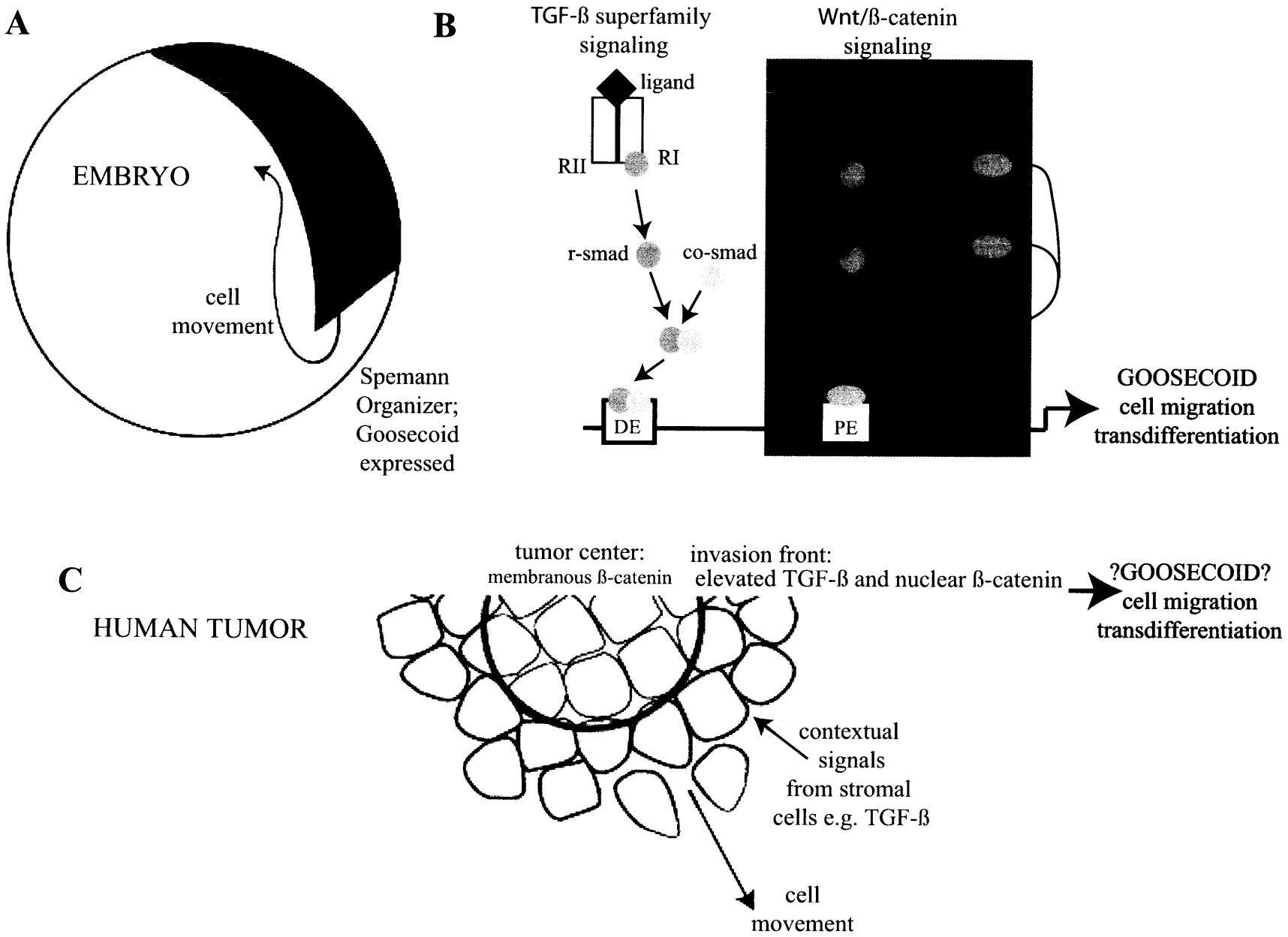
markers:
vimentin
fibronectin
Fsp-1
N-cadherin
MMPs
smooth-muscle actin

Figure 2. A hypothesized role for Goosecoid in tumor invasion

A. The induction of *Goosecoid* at the Spemann organizer is induced by the convergence of TGF-β and Wnt signals as shown.

B. The molecular mechanism by which the *Goosecoid* promoter is induced by these pathways is depicted in detail. Note that the induction of *Goosecoid* by β-catenin/tcf transcriptional activity is not direct.

C. Observations in the literature are compiled into a visual illustration of thesis hypothesis. Goosecoid may be induced in a tumor context where it recapitulates its functions in the organizer. One hypothesized mechanism for the induction of *Goosecoid* in a tumor setting is as a similar convergence of TGF-β and Wnt signals at the invasion front of tumors. Differential subcellular localization of β-catenin at the invasion front as well as differential levels of TGF-β at the invasion front suggested the possibility that *Goosecoid* might be upregulated at the tumor edge as a result of a convergence of these signals. The differential patterns of TGF-β and β-catenin across a tumor likely reflects the influence of stromal signals. PE, proximal element in *Goosecoid* promoter; DE, distal element in the promoter.


A
EMBRYO
cell
movement
Spemann
Organizer;
Goosecoid
expressed
B
TGF-ß superfamily
signaling
ligand
RII
RI
r-smad
co-smad
DE
Wnt/ß-catenin
signaling
PE
GOOSECOID
cell migration
transdifferentiation
C
HUMAN TUMOR
tumor center:
membranous ß-catenin
invasion front:
elevated TGF-ß and nuclear ß-catenin
?GOOSECOID?
cell migration
transdifferentiation
contextual
signals
from stromal
cells e.g. TGF-ß
cell
movement

Chapter 2

A Role for the Spemann Organizer Gene, *Goosecoid*, in Tumor Metastasis

Kimberly A. Hartwell[1,2], Jing Yang[1], Sendurai A. Mani[1], Sarah Frew[2], Lei Xu[2], Beth Muir[3], Ferenc Reinhardt[1], Anne E. Carpenter[1], George Bell[1], Mary Rockas[1], Dennis C. Sgroi[3], and Robert A. Weinberg[1,2]

[1]Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142 USA
[2]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA
[3]Department of Pathology, Harvard Medical School, Molecular Pathology Research Unit, Massachusetts General Hospital, Boston, Massachusetts 02129 USA

Published in abridged form in:
Hartwell KA, Muir B, Reinhardt F, Carpenter AE, Sgroi DC, Weinberg RA. The Spemann Organizer Gene, *Goosecoid*, Promotes Tumor Metastasis. Proc Natl Acad Sci USA. 2006 Dec 12;103(50):18969-74.

Weinberg R A (2007) *The Biology of Cancer* (Garland Science, New York).

FOXC2 collaboration (Figure 9B) published in:
Mani AS, Yang J, Brooks M, Schwaninger G, Zhou1 A, Miura N, Kutok JL, **Hartwell KA**, Richardson AL, Weinberg RA. Mesenchyme Forkhead 1 (FOXC2) plays a key role in metastasis and is associated with aggressive basal-like breast cancers. In review.

The analyses involving Twist were completed in collaboration with Jing Yang, and the analyses involving FOXC2 were completed in collaboration with Sendurai Mani. Lei Xu generated the cDNA samples of the increasingly aggressive metastatic melanomas and Sarah Frew performed the immunoblotting of IQGAP1 in Goosecoid-expressing HMECs. Dennis Sgroi participated in the design of the ductal-type clinical tumor analysis and in data analysis, Beth Muir performed the quantitative real-time RT-PCR using the ductal samples. Anne Carpenter leveraged the power of CellProfiler software to quantify the metastatic efficiency of the MDA-MB-231 cells in vivo. Ferenc Reinhardt assisted with animal husbandry and mouse injections, George Bell performed the bioinformatics analysis, and Mary Rockas assisted with multiple aspects of various projects, including the initial observation of SIP1 induction by Goosecoid. All other experimentation was completed by the thesis author, Kimberly Hartwell.

Elevated Goosecoid expression in human breast tumors

The expression patterns of the *Goosecoid* gene have not been well characterized in human or murine adult tissues. To determine whether a role for the *GSC* developmental gene in cancer was plausible, I undertook to examine human tumor specimens for evidence of *GSC* mRNA. Because probes for this gene were not included in published microarray expression studies to the best of my knowledge, I was unable to assess *GSC* expression patterns through database mining. We therefore measured *GSC* levels in a cohort of microdissected human breast tumors of three prevalent pathological subtypes: atypical ductal hyperplasia (ADH), ductal carcinoma *in situ* (DCIS), and invasive ductal carcinoma (IDC) (Ma et al. 2003). The 72 tumor samples examined were each accompanied by a patient-matched sample of normal breast epithelium. The normal samples were presumably proliferative per published studies of normal human breast tissue (Going et al. 1988). Because all samples in this cohort were obtained by laser capture microdissection, the samples do not contain significant numbers of stromal cells.

Quantitative real-time RT-PCR was used to compare levels of *GSC* mRNA represented in individual samples. The abundance of *GSC* mRNA in the normal tissue samples was found to be low, as signals were not detected until a high cycle number during PCR amplification. Strikingly, *GSC* expression was elevated in 56 out of 72 tumors (78%) compared to corresponding patient-matched normal tissue samples (Figure 1). By subtype, 71% of ADH samples, 79% of DCIS samples, and 78% of IDC samples contained a level of *GSC* mRNA above that of patient-matched normal tissue, and this pattern of *GSC* upregulation was found to be significant in each case ($p = 0.02$ for ADH, $p < 0.01$ for DCIS, $p = 0.01$ for IDC samples). The average extent of elevation of *GSC* mRNA across all samples per subtype was 5.9, 9.6, and 6.9 fold in the ADH,

DCIS, IDC samples, respectively, compared to corresponding normal samples. A more detailed view of the *GSC* expression data set can be found in Supplementary Table 1. We did not find a correlation between the expression level of GSC and the various clinical parameters accompanying these samples e.g., tumor grade, lymph node status, etc. We also performed nearest neighbor analysis in the interest of identifying genes that are functionally relevant to *Goosecoid* expression and Goosecoid function. More specifically, the observed *Goosecoid* expression pattern was compared to gene expression patterns previously recorded using this same sample set (Ma et al. 2003). The genes whose previously-assessed expression patterns correlated the most closely with *Goosecoid* were identified by statistical analysis. Genes that correlate closely (either positively or negatively) with *Goosecoid* are more likely to be relevant to *Goosecoid* expression and/or function than genes that do not. The resulting gene list obtained by this analysis did not contain anything that seemed especially mechanistically telling.

In summary, these results show that in a majority of human ductal-type breast tumors, *GSC* expression is significantly elevated above normal levels, consistent with a role for this developmental gene in human cancer, as hypothesized.

The generation of a Goosecoid-specific antibody

I generated a Gsc-specific antibody in order to assess *Gsc* activity at the protein level. Rabbit polyclonal antisera against Gsc were commercially produced by Covance using a KLH-conjugated peptide of the sequence CSENAEKWNKTSSSKA (common to both human and mouse Gsc). The specificity of the antisera was assessed by western immunoblotting using whole

cell lysates expressing either tagged or untagged ectopic Gsc (Figure 2), and the resulting antisera with confirmed specificity for Goosecoid protein were affinity purified.

I was especially interested in using this antibody to assess the protein expression of Goosecoid in clinical tumors *in situ*, given that I had hypothesized that Goosecoid might be induced by contextual signals at the invasion front of tumors. In an attempt to detect Goosecoid protein *in situ*, I tested the utility of this antibody for immunohistochemical staining. Primary tumors generated in mice by the injection of cells expressing either Gsc or control GFP protein were used as positive and negative control tissues. I confirmed by western blotting that the positive control tissues did indeed express significant quantities of Gsc (data not shown). To my great disappointment, I found that this Gsc antibody was not able to specifically detect Gsc on either frozen or formaldehyde-fixed, paraffin-embedded tissue sections (data not shown). Although I tested numerous variations of my immunoperoxidase protocol to no avail, it is possible that additional adjustments could render the antibody useful for this type of staining. I did observe, however, that the purified antiserum can be used to detect ectopic Gsc in formaldehyde-fixed cultured cells by immunofluorescence (data not shown).

Goosecoid elicits an epithelial-mesenchymal transition (EMT) and enhances cell motility

To identify the functional consequences of Gsc expression in adult epithelial cells, I stably expressed this protein in immortalized human mammary epithelial cells (HMECs) and in Madin-Darby canine kidney epithelial (MDCK) cells using retroviral transduction (Figure 3A). Neither of these parental cell lines expressed substantial levels of Gsc protein by western blotting

(Figure 3A). I noted that ectopic Gsc was localized to the nucleus upon examination in the HMECs (data not shown).

In both cell types, I observed that the population of cells expressing ectopic Gsc lost cell-cell contacts and displayed a scattered distribution in culture (Figure 3B), while control cells retained their typical epithelial morphology, continuing to grow as groups of cobblestone-like cells. The morphological changes evident in the Gsc-expressing cells were suggestive of an epithelial-mesenchymal transition (EMT). I therefore examined the status of known EMT markers in these cells. The Gsc-expressing cells demonstrated marked downregulation of E-cadherin, α-catenin and γ-catenin proteins, concordant with the apparent loss of adherens junctions (Figure 3C and 3D). These cells had replaced their cytokeratin-based intermediate filament network with one based on vimentin and stained positively for the mesenchymal protein N-cadherin (Figure 3C and 3D). The levels of the ECM component fibronectin also increased in these cells (albeit only minimally in the HMECs), and β-catenin levels dropped dramatically in the MDCK cells but not in the HMECs, which could be due to a possible defect in the β-catenin degradation pathway in these cells (data not shown) (Kikuchi 2000).

Moreover, the Gsc-expressing human mammary epithelial cells were found to be substantially more migratory in transwell migration assays than control cells (Figure 3E). They were also more invasive, as assessed similarly using transwells containing a matrigel barrier to migration (data not shown). My results demonstrate that Gsc induces the central hallmarks of an EMT and cell invasiveness in adult mammalian epithelial cells, recapitulating cellular changes driving gastrulation in higher vertebrates.

Goosecoid affects cell growth in vitro

In culturing the Gsc-expressing human and canine epithelial cells (HMEC and MDCK), I noted that the rate at which the Gsc-expressing cell populations expanded was significantly different than that of the GFP-expressing controls. Growth curves confirmed that the population doubling time for the Gsc-expressing populations was twice as long as that of controls (Figure 4). Similar observations have been made with the developmental EMT-inducing transcription factor Snail. This protein is reported to reduce the proliferation rate of MDCK cells *in vitro*, even while it confers metastasis-associated phenotypes on such cells (Vega et al. 2004). A proliferation assay (e.g. BrdU incorporation) would be necessary to confirm that the difference in the doubling time of the Gsc-expressing cell populations was due to a decrease in cell proliferation rather than a difference in cell survival rate. Gsc may, in fact, confer protection against apoptosis, as Snail is known to do. Snail has been shown to confer resistance to cell death in MDCK cells as assessed by a number of apoptosis assays (Vega et al. 2004).

TGF-β signaling induces Goosecoid expression in adult breast epithelial cells

The Wnt/β-catenin and TGF-β superfamily signaling cascades are required for Spemann organizer formation and *Gsc* gene expression (Harland and Gerhart 1997) and these same pathways have been implicated in tumor metastasis (Thiery 2002). Since Gsc recapitulated aspects of its embryonic organizer function in adult mammalian epithelial cells, I tested whether these two organizer-associated signaling cascades induce *GSC* expression in these cells. I found

that the enhancement of Wnt/β-catenin signaling by two approaches failed to activate *GSC* expression. Specifically, *GSC* mRNA expression was not increased in human mammary epithelial cells (HMECs), either by expression of a non-degradable form of β-catenin (ΔN90 β-catenin (Barth et al. 1997)) or by a constitutively-active form of Lef-1 (Lef-vp16 (Aoki et al. 1999)), a DNA-binding protein that associates with β-catenin to induce transcription of target genes (Figure 5A). I confirmed these constructs were transcriptionally functional using the Topflash/Fopflash reporter system (data not shown) (Korinek et al. 1997).

In contrast, expression of constitutively-active TGF-β type 1 receptor (Wieser et al. 1995) in these cells using retroviral transduction did induce *GSC* mRNA expression (Figure 5B). *GSC* mRNA was also induced in non-transduced HMECs in a dose-dependent manner by the addition of soluble, activated TGF-β1 to the cell culture medium (Figure 5C). When TGF-β1 was applied to HMECs expressing non-degradable β-catenin or the constitutively-active form of Lef-1 or green fluorescent protein (GFP) control, *GSC* expression was not induced to a level greater than that achieved without activation of the Wnt/β-catenin pathway (data not shown).

Together, these experiments demonstrate that TGF-β signaling induces *GSC* in adult breast epithelial cells as do related mesoderm-inducing signals in gastrulating embryos and other cells (Watabe et al. 1995; Labbe et al. 1998; Ku et al. 2005). I note here that the induction of *GSC* observed in HMECs was not extensive, suggesting that a robust induction of *GSC* in these cells requires cooperative signaling, as in *Xenopus* embryos. I did not observe β-catenin acting synergistically with these signals in our HMEC system, contrary to observations in *Xenopus* embryos (Watabe et al. 1995). This may reflect distinct roles for Goosecoid in *Xenopus* and humans, as well as distinct mechanisms of regulation.

The induction of Goosecoid by TGF-β signaling coincides with an EMT

The described induction of *GSC* by TGF-β signaling in the human mammary epithelial cells (HMECs) coincided with the induction of an EMT. This EMT was evidenced by the suppression of epithelial E-cadherin and induction of the mesenchymal components vimentin and fibronectin, as judged by western immunoblotting (data not shown). As previously discussed, TGF-β signaling is known to be capable of generating an EMT in mammalian epithelial cells in the presence of Ras and other signals (Grunert et al. 2003). I do not yet know whether *GSC* is essential for the observed TGF-β-induced EMT. One way to address this question is to antagonize *GSC* expression using short interfering RNA (siRNA)-mediated inhibition (Elbashir et al. 2001). To this end, I generated eight lentiviral shRNA-expressing constructs directed against human *GSC*. By quantitative real-time RT-PCR, I observed that only one out of the eight constructs reproducibly generated a significant decrease in *GSC* mRNA transcript. In each of the two breast cancer cell lines tested, namely MCF7-Ras and Sum1315, endogenous *GSC* was reduced by at least 80% compared to its original levels (data not shown). This construct will hopefully prove useful in addressing the question of whether *GSC* mediates certain TGF-β-induced cellular effects, and in addressing other biological questions.

The identification of additional genes that may regulate Goosecoid

To identify genes that might be capable of inducing *GSC* within a human tumor, in addition to TGF-β, I employed an alternative approach. I used bioinformatics software to search

for putative binding sites within the 10kb region of the human, mouse, and rat genomes that is immediately upstream of the *Goosecoid* coding region. By performing a combined search with these three promoters, I hoped to identify well conserved genes, increasing my chances of identifying meaningful candidates.

A subset of the putative binding sites identified in this way is shown in Table 1. All of the binding sites listed are conserved across the three species. Intriguingly, several of the candidate transcription factors identified are components of pathways that are already tied to the EMT and cell motility, as detailed in Table 1. For example, NF-κB, a component of the TNF-α pathway, was identified, as was Oct-1, and these two transcription factors can cooperatively regulate target genes (Voleti and Agrawal 2005). Moreover, NF-κB was recently found to be essential for the induction of an EMT in mammary epithelial cells by TGF-β and for the metastasis of these cells *in vivo* (Huber et al. 2004). I am thus currently testing whether TNF-α/NF-κB signaling is able to regulate the expression of *GSC* in human cells.

Goosecoid expression patterns in cancer cell lines

My observation of a Gsc-induced EMT in human and canine epithelial cells prompted me to examine whether *GSC* expression correlates with the EMT status of established cancer cell lines. To this end, I performed RT-PCR to assess the levels of *GSC* mRNA and *GSC*-like mRNA in a large number of cancer cell lines.

I first examined the expression levels of these genes in 15 breast cancer cell lines and found that, while *GSC* was detectable in several of these lines, its expression did not consistently

correlate with the EMT phenotype or with the known aggressiveness of these lines (Figure 6A). *GSC*-like expression was not detected in any of these lines (data not shown).

To assess the levels of *GSC* expression in cancer cell lines from other tumor types, I applied for and received a set of 59 different cancer cell line samples from the National Cancer Institute. This set is known as the NCI-60 panel of human cancer cell lines. I again found that the expression levels of *GSC* did not correlate with EMT status and cell line aggressiveness (Figure 6B). In fact, *GSC* was only weakly detected in a few of these cell lines after a full 30 PCR cycles.

In addition to these analyses, we also examined the levels of *GSC* expression in a set of increasingly metastatic melanoma cell lines in collaboration with Dr. Lei Xu in the lab of Richard Hynes at MIT. This set includes four highly metastatic human cell lines, MA-1, MA-2, MC-1, MC-2, which Dr. Xu derived from the weakly metastatic melanoma cell line A375P by the *in vivo* selection of progressively metastatic variants as described (Clark et al. 2000; Xu et al. 2006). *In vivo*, by the experimental metastasis assay, these five lines are increasingly metastatic, as indicted by the number of lung metastases formed, in the following order: A375P (least metastatic), MA-1, MA-2, MC-1, MC-2 (most metastatic). Intriguingly, we observed that the expression level of *GSC* in primary tumors generated from the subcutaneous injection of these variants tightly correlated with the known metastatic ability of the lines (Figure 6C). This observation suggests that GSC expression promotes the metastatic phenotype in melanoma. The question of whether or not GSC is required for the *in vivo* metastatic behavior of these aggressive derivative cell lines has not yet been decisively answered.

Notably, this provocative pattern of *GSC* expression in these A375P derivative melanoma lines was not mirrored in the melanoma lines that are included in the NCI-60 panel, which did

not show evidence of *GSC* expression. This difference may reflect the fact that the cells in the NCI-60 panel were growing as two-dimensional *in vitro* cultures at the time of sampling, whereas the A375 derivative lines were growing *in vivo* as primary tumors. The A375 derivative lines had also not been extensively cultured prior to their injection into mice. Therefore, one hypothesis that might explain this inconsistency is that while *GSC* may be active in melanomas *in vivo*, it is selected against during the *in vitro* culturing of melanoma cells. Alternatively, *GSC* expression may require the continued exposure to contextual signals that are present in an *in vivo* context but that are absent *in vitro* cultures.

Goosecoid enhances the metastatic ability of cancer cells

Because Gsc triggered an EMT and enhanced cell motility in adult epithelial cells- both known correlates of invasive and metastatic ability - I tested whether this gene could also promote tumor metastasis. Gsc was ectopically expressed in GFP-labeled MDA-MB-231 human breast cancer cells (Figure 7A). The cells of this line are weakly metastatic and quasi-mesenchymal, in that they do not express E-cadherin and do express vimentin, yet they display an epithelial-like morphology in culture (Price et al. 1990). I observed that upon the introduction of Gsc, the MDA-MB-231 cells acquired a spindle-like morphology more typical of mesenchymal cells (Figure 7B) as well as an increased degree of motility (Figure 7C).

Control or Gsc-expressing MDA-MB-231 cells were injected into the tail veins of mice and lungs were examined for metastases eight to ten weeks after injection (Figure 7D). At both timepoints, a greater number of pulmonary metastases were visible in the mice injected with

Gsc-expressing cells. Quantification of the observed lung nodules at eight weeks using image analysis indicated a four-fold increase in the average number of metastases in the mice injected with Gsc-expressing cells compared to control animals (Figure 7E).

This demonstrated enhancement of metastasis might have arisen as a consequence of a Gsc-induced stimulation of proliferation *in vivo*. To address this possibility, I directly compared the *in vivo* proliferation rates of these two cell populations by injecting them either into the subcutaneous space or into the mammary glands of mice. In fact, the resulting primary tumors generated by the Gsc-expressing cells grew more slowly than did control tumors at both sites (Figure 8 and data not shown).

The demonstration that Gsc-expressing breast cancer cells formed significantly greater numbers of metastases in murine lungs in spite of proliferating more slowly *in vivo* provide strong indication that Gsc expression enhances the metastatic ability of MDA-MB-231 human breast cancer cells.

I note here that other carcinoma cell lines were similarly tested *in vivo*, including MCF-7 (derived from human breast cancer), MCF-7 Ras, HAIER (human kidney), NuMG (mouse mammary), HMLER (human mammary), and Sum 149 (human breast). Preliminary experiments did not reveal a Gsc-induced enhancement of metastasis in any of these other lines (data not shown).

Goosecoid may silence E-cadherin by inducing SIP1

One important question that remains is how Gsc mediates its cellular effects, including the induction of an EMT. As previously mentioned, other developmental transcription factors have been linked to the EMT and metastasis during the course of this work, namely SNAI1 (Snail), SNAI2 (Slug), SIP1(ZEB2), Twist, δEF1 (ZEB1), and E47 (E12). All have been found to transcriptionally repress *E-cadherin* by directly binding the E-box elements within the proximal promoter (Hajra et al. 2002) (Eger et al. 2005) (Perez-Moreno et al. 2001; Yang et al. 2004). I likewise reasoned that Gsc might directly bind and silence the *E-cadherin* promoter. However, extensive analysis of the DNA sequence of the human *E-cadherin* promoter region, including the gene coding region plus 10kb of flanking sequence on both ends (which contains the E-boxes), revealed a surprising absence of a putative Gsc binding site. It may be, therefore, that Gsc is an anomaly among these developmental transcription factors in that it does not induce an EMT by regulating *E-cadherin* directly. The use of a reporter construct under the control of the *E-cadherin* promoter would help to clarify this question.

Alternatively, Gsc might silence the expression of E-cadherin indirectly, by binding instead to one of the six aforementioned developmental transcription factors. By quantitative real-time RT-PCR, I observed a substantial induction of *SIP1* mRNA by Gsc in HMECs, and only minimal induction of *Snail* mRNA (Figure 9A). In contrast, *Slug* mRNA was not increased in the Gsc-expressing cells compared to controls. In collaboration with Dr. Jing Yang in our lab, we also found that Twist was not induced by Gsc expression, as assessed by western immunoblotting (data not shown). These observations suggest that Gsc silences E-cadherin by

selectively inducing *SIP1*. Indeed, SIP1 may be required for a Gsc-induced EMT, but this possibility has not yet been addressed.

Goosecoid induces the expression of FOXC2

Another transcription factor that has been linked to metastasis, specifically to the process of metastatic colonization, is Foxc2 (Yang et al. 2004). Foxc2 is expressed in cells of the developing mouse that are destined to form mesoderm (Sasaki and Hogan 1993). As previously discussed, a screen for genes whose expression patterns correlated with the cellular ability to bypass distinct barriers to metastasis *in vivo* was completed in our laboratory. In this analysis, *Foxc2* was identified as a possible colonization gene, as it was expressed only in the derivative cell line that could successfully colonize the lungs (Yang et al. 2004). Further characterization of *FOXC2* in our lab has confirmed a functional role for this gene in metastasis (unpublished results, S. A. Mani et al., manuscript in preparation). In collaboration with Dr. Sendurai Mani in our lab, FOXC2 was found to be induced by Gsc expression in HMECs by western immunoblotting (Figure 9B). We do not yet know whether this induction of FOXC2 is essential to Gsc function. Notably, I also tested whether *Gsc* is expressed in the cell lines comprising the metastasis model in which Foxc2 was linked to colonization (Yang et al. 2004). Again, this model consists of four cell lines that were all derived from a single mouse mammary tumor. Upon re-injection into the orthotopic site, these lines generate primary tumors with equivalent growth kinetics but distinct metastatic behaviors. I found that *Gsc* is not expressed in any of

these four cell lines, suggesting that while Gsc may be an important inducer of *FOXC2* in tumors, it is unlikely to be the only means by which this gene can become active (data not shown).

Goosecoid most likely does not induce an EMT via a TGF-β autocrine loop

It has been reported in the literature that continuous TGF-β signaling is required for the maintenance of an EMT in certain settings (Grunert et al. 2003). I therefore reasoned that another possible mechanism by which Gsc might induce an EMT is by the induction of such a TGF-β autocrine loop. In addressing this question, I employed a TGF-β antibody and a chemical inhibitor from Biogen IDEC. I also used a commercially available antibody that antagonizes TGF-β signaling by binding to its ligands.

In the presence of these different inhibitors, the Gsc-expressing and control populations of HMECs were re-derived, in that ectopic Gsc was expressed in a population of HMECs that had not previously expressed this gene. Despite the presence of the various inhibitors, Gsc was able to induce a robust EMT in these cells under all treatments (data not shown). This observation is preliminary, as the complete antagonism of TGF-β signaling by these inhibitors during the course of this experiment has not yet been confirmed. Other members of the Weinberg lab have confirmed that these inhibitors are functional in HMEC cells, however. If the inhibitors are found to have successfully prevented any induction of TGF-β signaling in these experiments, I would conclude that Gsc most likely does not trigger an EMT by generating a TGF-β autocrine loop.

A bioinformatics approach identifies IQGAP1 as downstream of Goosecoid

I reasoned that an unbiased search for the presence of putative Gsc binding sites across the promoters of known genes might assist in elucidating signaling downstream to Gsc. I therefore again took a bioinformatics approach and screened the DNA sequence of a collection of known human gene promoters compiled by scientists at the Whitehead Institute. A subset of the gene promoters containing putative Gsc binding sites is shown in Table 2.

Again, many of the genes identified in this way have already been linked to the EMT and/or metastasis. One gene in particular stood out on this list, IQGAP1, because it had also been reported to correlate with the metastatic phenotype in increasingly metastatic cell lines derived from the A375 melanoma parental line as described (Clark et al. 2000). IQGAP1 is a scaffolding protein that has known roles in the regulation of the cytoskeleton, cell adhesion, and GTPase signaling (Briggs and Sacks 2003). Because we had observed that *Gsc* expression also correlated with the metastatic capability of A375 derivative lines, the presence of a putative Gsc binding site in the IQGAP1 human promoter suggested that this gene might in fact be a direct transcriptional target of Gsc in a tumors. In collaboration, Sarah Frew in the Hynes lab at MIT assessed the levels of IQGAP1 in the Gsc-expressing HMECs by western immunoblotting. We observed that Gsc does upregulate the expression of IQGAP1 in the human epithelial HMEC cells by 2.6 fold (Figure 9C), indicating that Gsc might indeed function through IQGAP1. Such a mechanism could be active both in certain melanomas and carcinomas, given the nature of our observations.

PDGFR signaling may mediate Goosecoid function

The PDGF signaling cascade is known to be essential for *Xenopus* gastrulation (Ataliotis et al. 1995), and has fundamental mitogenic and chemoattractive functions in mesenchymal cells (Elenbaas and Weinberg 2001). Intriguingly, I found that both the PDGF β-type receptor and PDGF β-type ligand were induced in the Gsc cells that had undergone EMT (Figure 9D and 9E). These observations were made by immunoblotting and quantitative real-time RT-PCR, respectively. The acquisition of a PDGF autocrine loop by way of Gsc induction might allow tumor cells to become more invasive (Jechlinger et al. 2003). At this time, I have not managed to consistently demonstrate a function for this upregulation of PDGF pathway components in the Gsc-expressing HMECs, despite intensive experimental efforts using both PDGFR signaling agonists (PDGF-α and -β ligands) and antagonists (Glivec) and a variety of assays (data not shown).

Figure 1. Quantification of *Goosecoid* expression in human tumors

The relative level of *GSC* mRNA in each tumor (blue) and corresponding normal (red) tissue sample is shown with the lowest value of each pair in foreground. Pairs are grouped by tumor pathological subtype and sorted within groups according to the level of *GSC* mRNA in the tumor samples. All values displayed were normalized to the average of the *GSC* mRNA levels in the normal samples, which is set as the y-value 1 in the graph. Values outside the scale of the y-axis are marked by an asterisk.

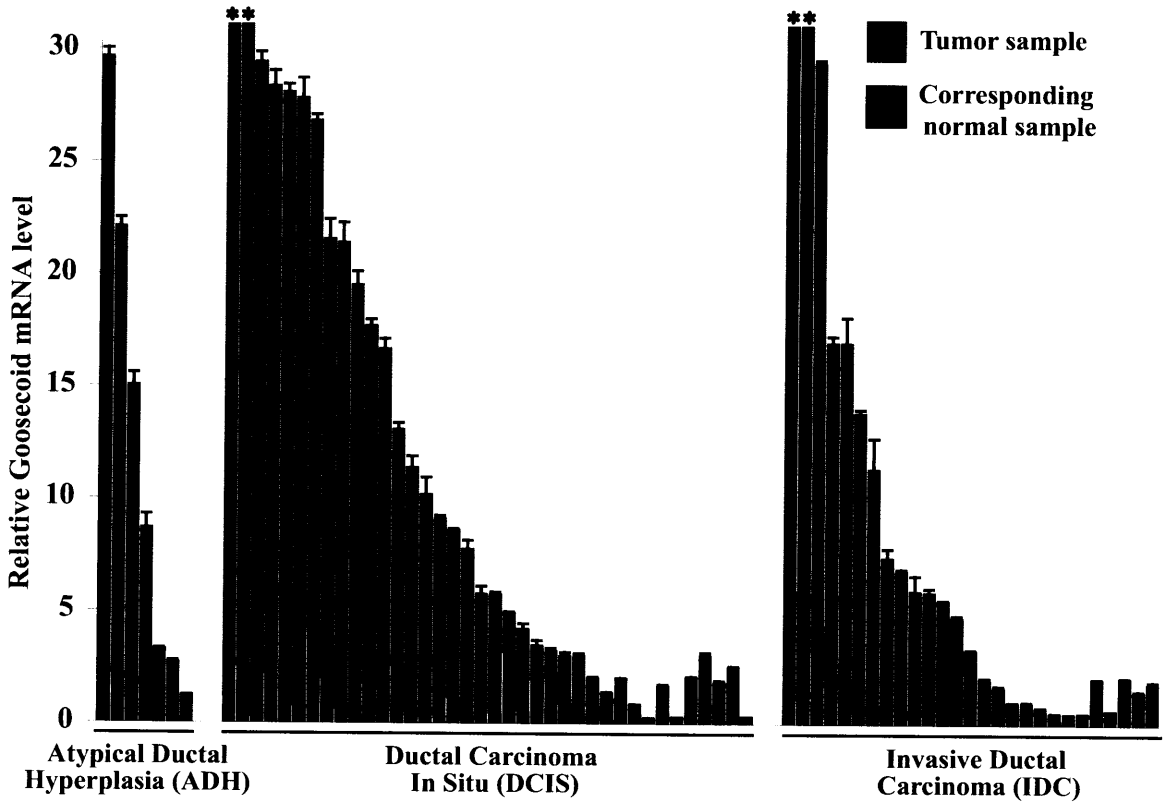

Relative Goosecoid mRNA level
30
25
20
15
10
5
0
**
**
Tumor sample
Corresponding normal sample
Atypical Ductal Hyperplasia (ADH)
Ductal Carcinoma In Situ (DCIS)
Invasive Ductal Carcinoma (IDC)

Figure 2. The generation of a Goosecoid antibody

The specificity of rabbit polyclonal antiserum generated against Goosecoid is confirmed by western immunoblotting. Untagged Goosecoid is detected at 33kD, as expected, in lysates of human mammary epithelial cells into which ectopic untagged Goosecoid was transduced. Tagged Goosecoid runs slightly slower, in lysates of human mammary epithelial cells that were instead transduced with ectopic tagged Goosecoid.

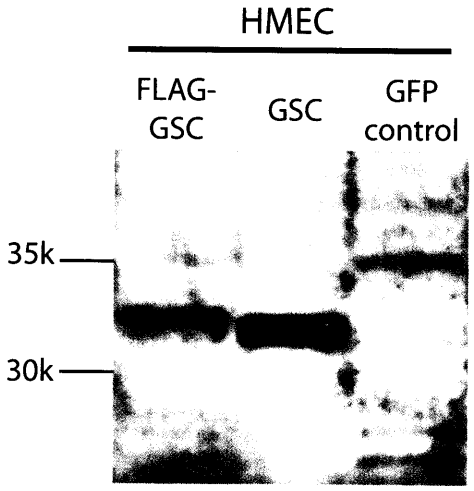

HMEC
FLAG-GSC
GSC
GFP control
35k
30k

Figure 3. Effects of Goosecoid expression in immortalized human breast and canine kidney epithelial cells

A. Ectopic expression of Gsc in human mammary epithelial cells (HMECs) and in Madin-Darby canine kidney (MDCK) epithelial cells, by western blotting.

B. Phase-contrast micrographs of HMECs and MDCK cells expressing either Gsc or GFP control.

C. Expression levels of epithelial proteins E-cadherin, α-catenin, and γ-catenin, and mesenchymal proteins N-cadherin and vimentin in HMECs and MDCK cells expressing either Gsc or GFP control, by western blotting. β-actin protein is shown as a loading control.

D. Immunofluorescence staining for epithelial proteins E-cadherin and cytokeratins, and mesenchymal protein vimentin in MDCK cells expressing either Gsc or GFP control. Antibody staining is shown in red, Hoechst nuclear staining in blue.

E. Quantification of the migratory abilities of HMECs expressing Gsc or GFP control by transwell migration assay. Movement toward medium with or without growth factor supplements (EGF, insulin and hydrocortisone) is graphed as the percent of total cells assayed that migrated after 48 hours. Assays were done in triplicate and the averages with sem are shown.

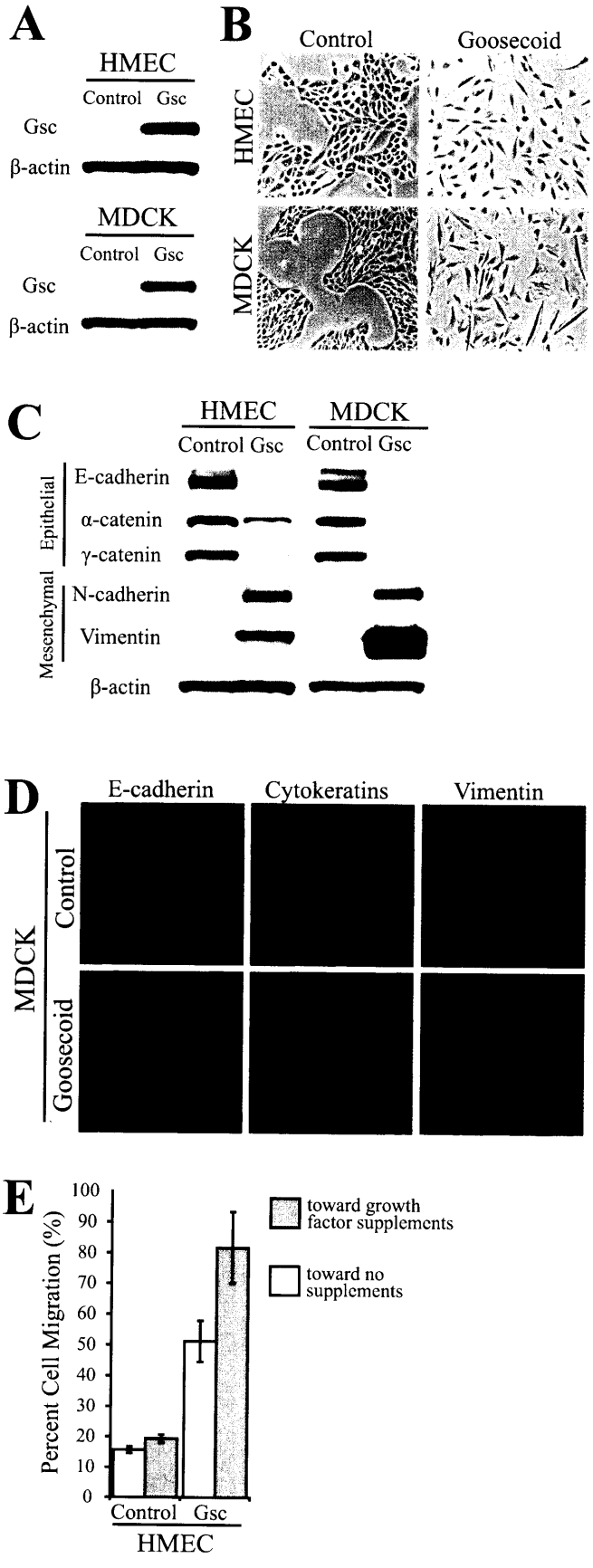
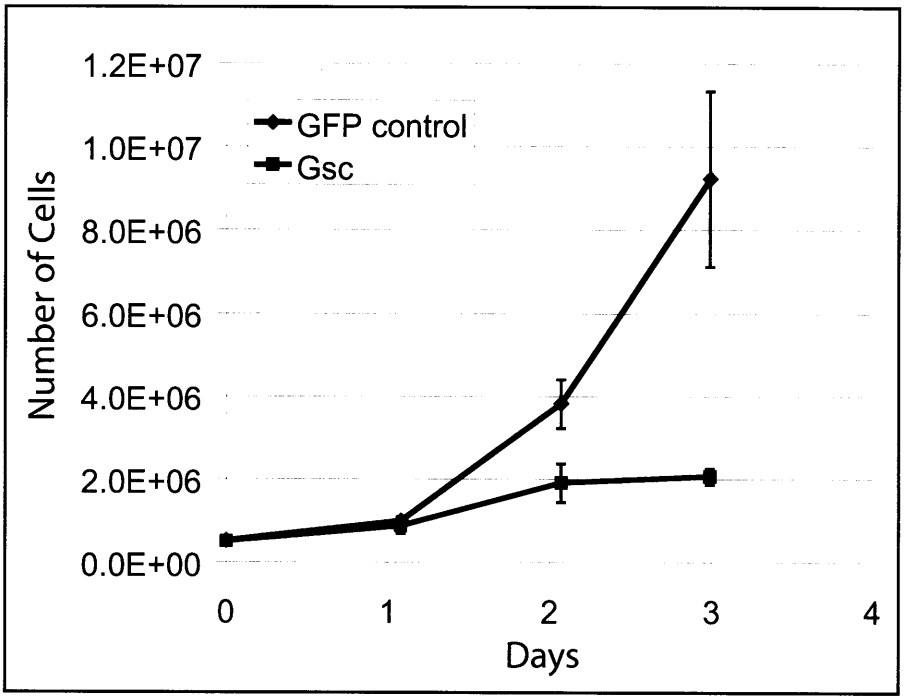

A
HMEC
Control Gsc
Gsc
β-actin
MDCK
Control Gsc
Gsc
β-actin
B
Control
Goosecoid
HMEC
MDCK
C
HMEC
Control Gsc
MDCK
Control Gsc
Epithelial
E-cadherin
α-catenin
γ-catenin
Mesenchymal
N-cadherin
Vimentin
β-actin
D
E-cadherin
Cytokeratins
Vimentin
MDCK
Control
Goosecoid
E
Percent Cell Migration (%)
100
90
80
70
60
50
40
30
20
10
0
toward growth factor supplements
toward no supplements
Control Gsc
HMEC

Figure 4. Effects of Goosecoid expression on population growth *in vitro*

The growth kinetics of polyclonal populations of Goosecoid-expressing human mammary epithelial cells (HMECs) and canine kidney epithelial cells (MDCKs) cultured *in vitro* are shown. Also shown are the kinetics of control GFP-expressing populations.

MDCK


1.2E+07
1.0E+07
8.0E+06
6.0E+06
4.0E+06
2.0E+06
0.0E+00
Number of Cells
GFP control
Gsc
0
1
2
3
4
Days

HMLE

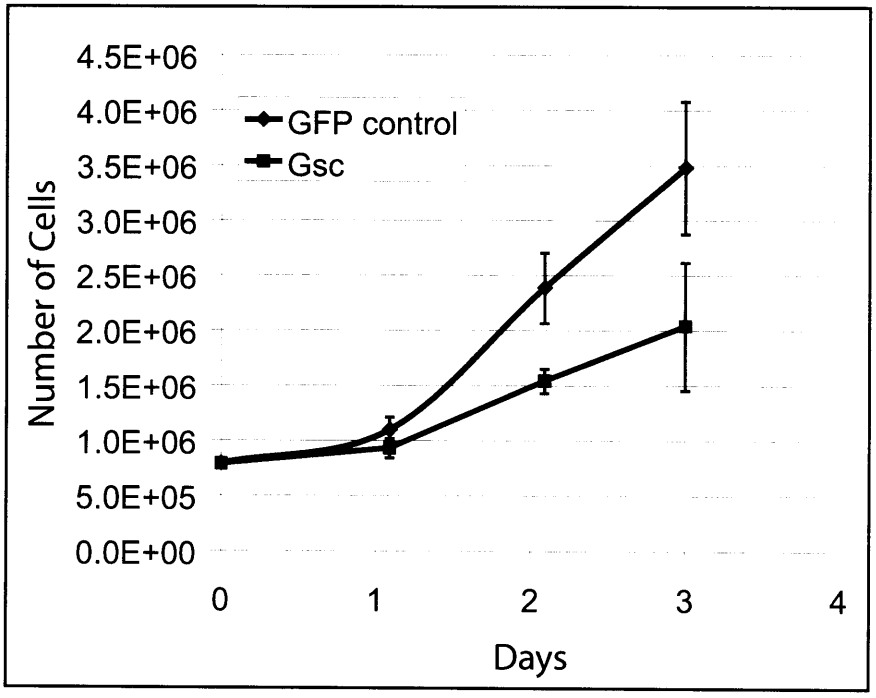

4.5E+06
4.0E+06
3.5E+06
3.0E+06
2.5E+06
2.0E+06
1.5E+06
1.0E+06
5.0E+05
0.0E+00
Number of Cells
GFP control
Gsc
0
1
2
3
4
Days

Figure 5. Induction of *Goosecoid* in human mammary epithelial cells (HMECs)

A. Relative *GSC* mRNA expression levels in HMECs containing empty vector, non-degradable β-catenin (ΔN90 β-cat), or constitutively active Lef-1 (Lef-vp16). Each bar represents the average with sem of triplicate assays.

B. Relative *GSC* mRNA expression levels in human mammary epithelial cells (HMECs) expressing either empty vector or constitutively-active TGF-β type 1 receptor. Each bar represents the average with sem of triplicate assays.

C. Relative *GSC* mRNA expression levels in HMECs treated with activated TGF-β1 ligand at various concentrations for three or six days. Each bar represents the average with sem of triplicate assays.

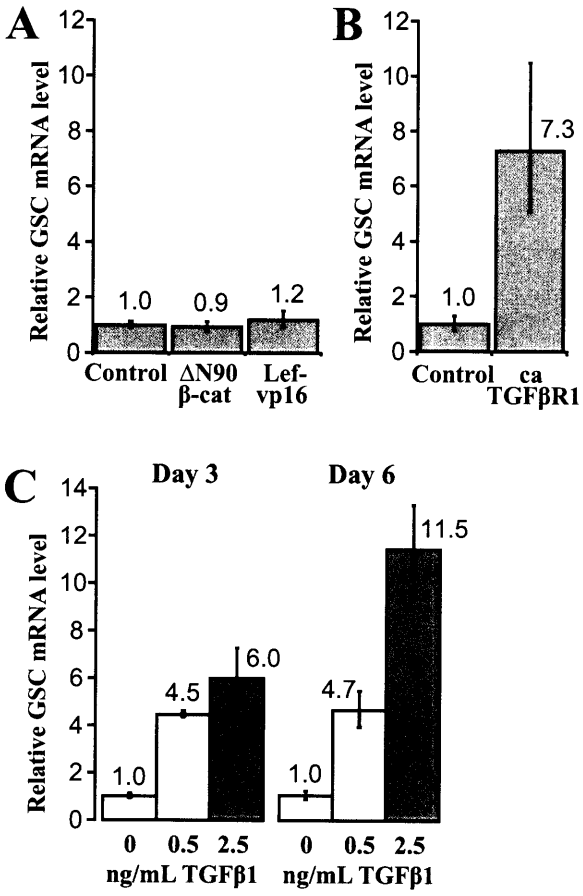

A
Relative GSC mRNA level
12
10
8
6
4
2
0
1.0
0.9
1.2
Control
ΔN90 β-cat
Lef-vp16
B
Relative GSC mRNA level
1.0
7.3
Control
ca TGFβR1
C
Day 3
Day 6
Relative GSC mRNA level
14
12
10
8
6
4
2
0
1.0
4.5
6.0
1.0
4.7
11.5
0
0.5
2.5
ng/mL TGFβ1
0
0.5
2.5
ng/mL TGFβ1

Table 1. Other putative signals upstream of *Goosecoid*

Listed in the table is a subset of the proteins found to have putative binding sites within the human, mouse, and rat *Goosecoid* gene promoters. All proteins listed had putative binding sites that were conserved across all three species. Pertinent information is also listed for each candidate.

PROTEIN	PATHWAY	KNOWN ROLES OF RELEVANCE	OTHER RELEVANT INFO
NFκB, nuclear factor κB	TNFα pathway	EMT and metastasis	cooperates with OCT1 (below)
OCT1, octamer-binding protein 1		embryonic morphogenesis and cell differentiation	cooperates with NFκB (above)
FOXD3, forkhead box D3		dorsal mesoderm formation	expression is organizer specific
SRF, serum response factor	numerous	cell cycle regulation, apoptosis, cell differentiation; EMT and invasion	
AP1, activator protein 1	numerous	EMT and invasion	fos and jun heterodimer
LEF1, lymphoid enhancer binding factor 1	Wnt pathway	EMT and invasion	beta-catenin binding partner
TCF4, transcription factor 4	Wnt pathway	EMT and invasion	beta-catenin binding partner

Figure 6. *Goosecoid* expression patterns in cultured transformed cell lines

A. *Goosecoid* expression levels are shown across a set of breast cancer cell lines by RT-PCR. Cell lines known to be especially aggressive are marked by a red asterisk. More benign lines are marked by a green asterisk.

B. GSC expression in the NCI-60 panel of 59 human transformed cell lines is shown, by tissue type. RT-PCR products are shown.

C. The expression pattern of GSC in a set of increasingly aggressive metastatic melanoma variant cell lines by RT-PCR is shown. Samples were generated from tumors that were generated by the subcutaneous injection of these lines into mice.

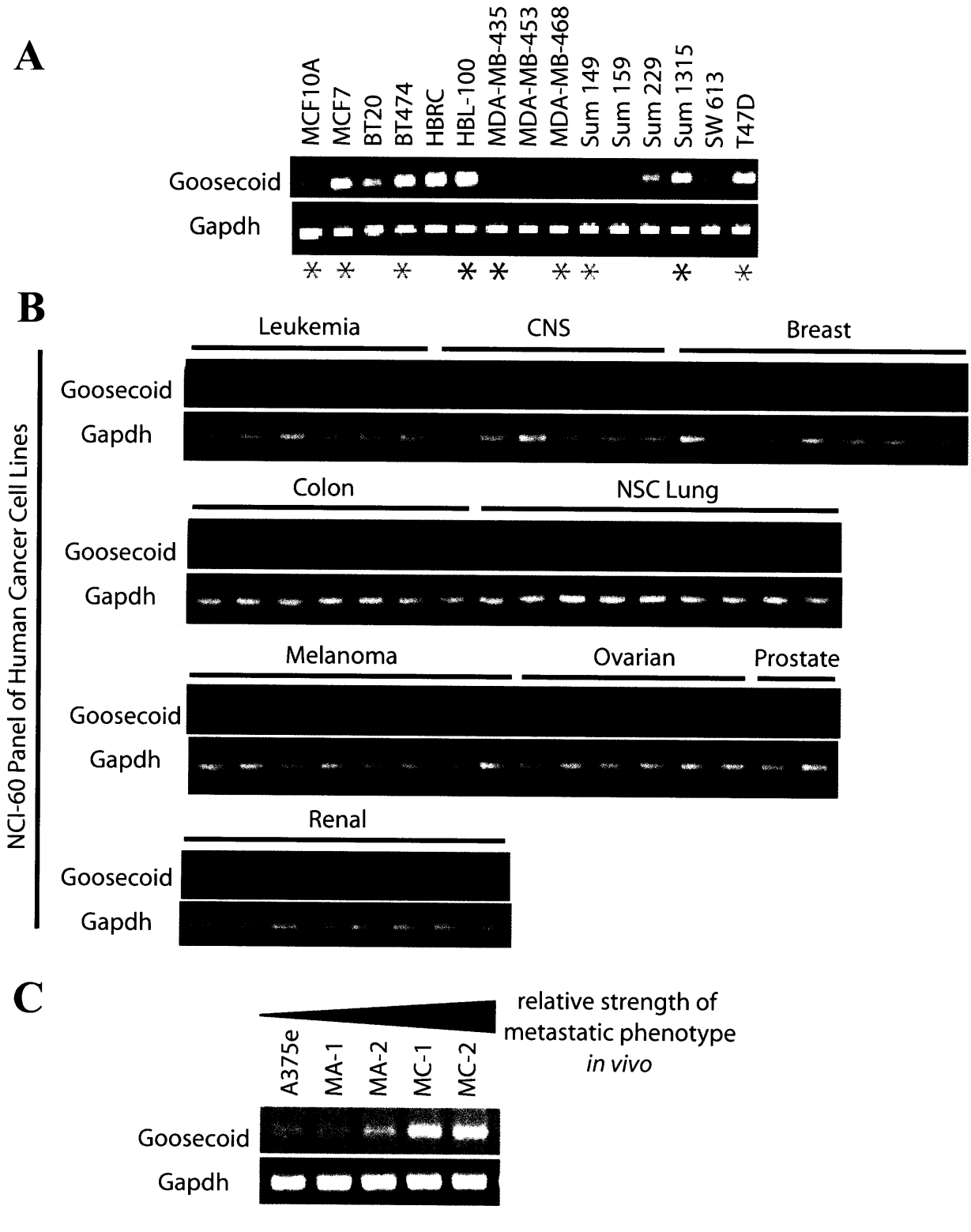

A
MCF10A
MCF7
BT20
BT474
HBRC
HBL-100
MDA-MB-435
MDA-MB-453
MDA-MB-468
Sum 149
Sum 159
Sum 229
Sum 1315
SW 613
T47D
Goosecoid
Gapdh
B
NCI-60 Panel of Human Cancer Cell Lines
Leukemia
CNS
Breast
Goosecoid
Gapdh
Colon
NSC Lung
Goosecoid
Gapdh
Melanoma
Ovarian
Prostate
Goosecoid
Gapdh
Renal
Goosecoid
Gapdh
C
relative strength of
metastatic phenotype
in vivo
A375e
MA-1
MA-2
MC-1
MC-2
Goosecoid
Gapdh

Figure 7. Goosecoid expression changes the behavior of MDA-MB-231 human breast cancer cells *in vitro* and in mice

A. Gsc expression in MDA-MB-231 cells expressing either Gsc or GFP control, by western blotting.

B. Phase-contrast micrographs of MDA-MB-231 cells expressing either Gsc or GFP control.

C. Quantification of the migratory abilities of MDA-MB-231 cells expressing Gsc or GFP control by transwell assay, graphed as the percent of total cells assayed that migrated after 16 hours. Assays were done in triplicate and the averages with sem are shown.

D. Representative brightfield and fluorescence images of mouse lung lobes 10 or 8 weeks post tail vein injection of MDA-MB-231 cells expressing either Gsc or GFP control.

E. Quantification of the number of metastatic foci in the lungs of mice 8 wks post tail vein injection of MDA-MB-231 cells expressing either Gsc or GFP control ($n \geq 6$, trend was confirmed by four independent experiments). Quartiles, medians, and the p value of the mean are shown.

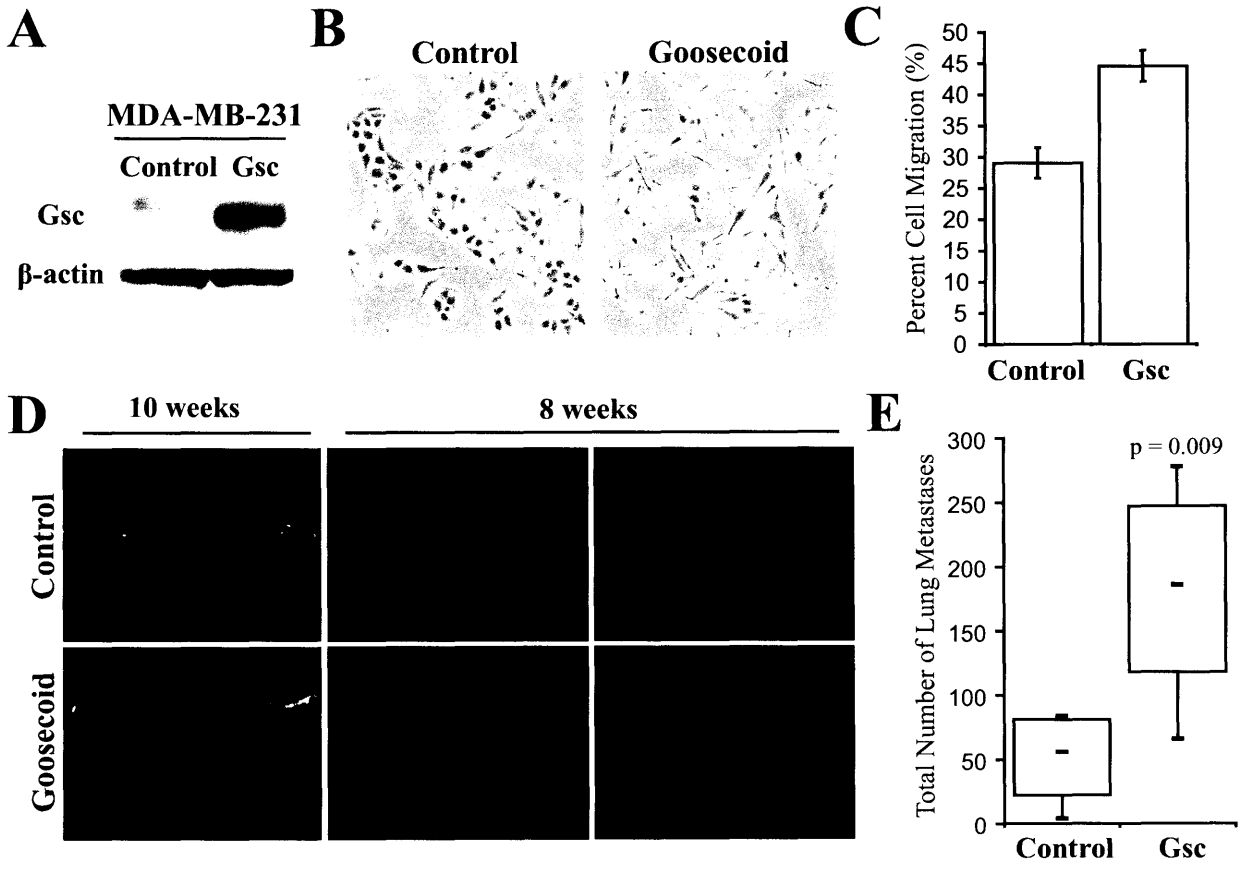

A
MDA-MB-231
Control Gsc
Gsc
β-actin
B
Control
Goosecoid
C
Percent Cell Migration (%)
50
45
40
35
30
25
20
15
10
5
0
Control
Gsc
D
10 weeks
8 weeks
Control
Goosecoid
E
Total Number of Lung Metastases
300
250
200
150
100
50
0
p = 0.009
Control
Gsc

Figure 8. Growth kinetics of primary MDA-MB-231 tumors

The growth kinetics of primary tumors generated by the subcutaneous injection of Gsc-expressing and control MDA-MB-231 cancer cell lines are shown.

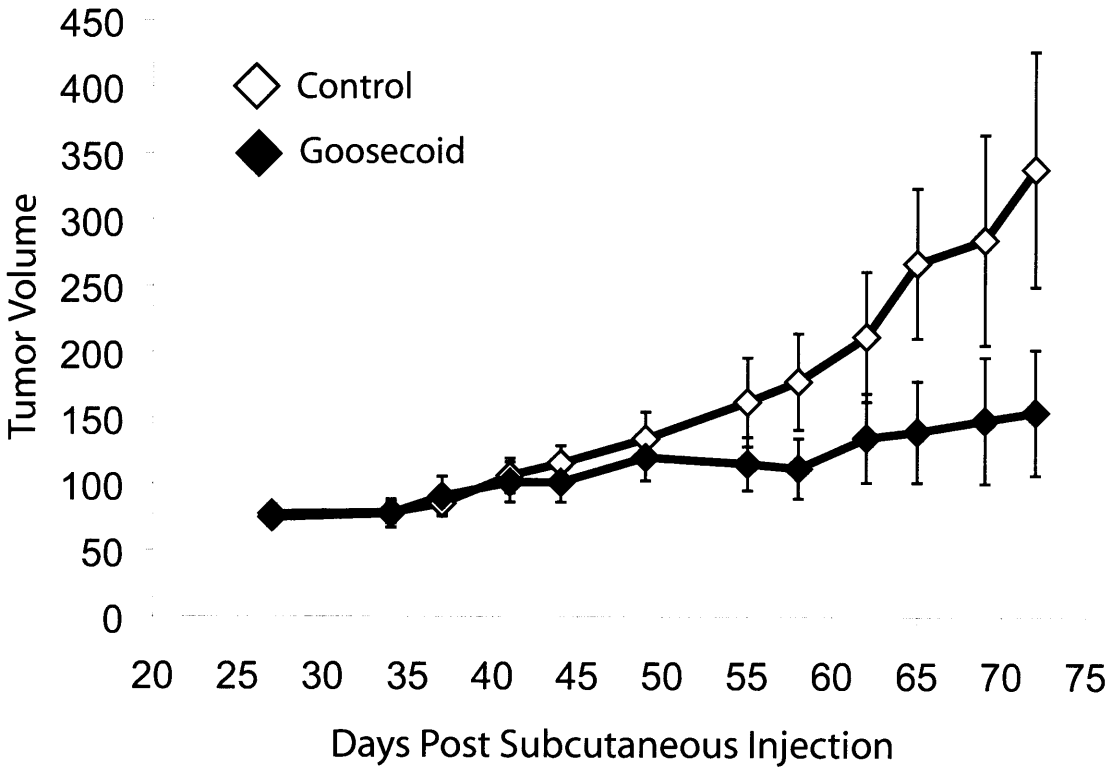

450
400
350
300
250
200
150
100
50
0
Tumor Volume
Control
Goosecoid
20
25
30
35
40
45
50
55
60
65
70
75
Days Post Subcutaneous Injection

Figure 9. The identification of additional Goosecoid target genes

A. Gsc robustly induces SIP1 expression by quantitative real-time RT-PCR in human mammary epithelial cells (HMECs). Gsc induces Snail only minimally and does not induce Slug. The 'low Gsc' cell population expresses only trace amounts of Gsc and is useful as an additional negative control. Fold induction is listed above the bars of the graph.

B. Gsc induces the expression of FOXC2 in HMECs by western immunoblotting. β-actin is shown as a loading control.

C. IQGAP1 is also induced by Gsc expression in HMECs. Western immunoblotting is shown is the fold upregulation in the Gsc expressing cells relative to control GFP cells. The 'low Gsc' cell population expresses only trace amounts of Gsc and is useful as an additional negative control. Nucleoporin is used as a loading control.

D. The PDGFR-β type receptor is upregulated at the protein level in HMECs, by western immunoblotting. The 'low Gsc' cell population expresses only trace amounts of Gsc and is useful as an additional negative control. β-actin is used as a loading control.

E. Likewise, the PDGF-β type ligand is upregulated in HMECs, as assessed by real-time quantitative RT-PCR, and is seen to correlate with *Goosecoid* levels.

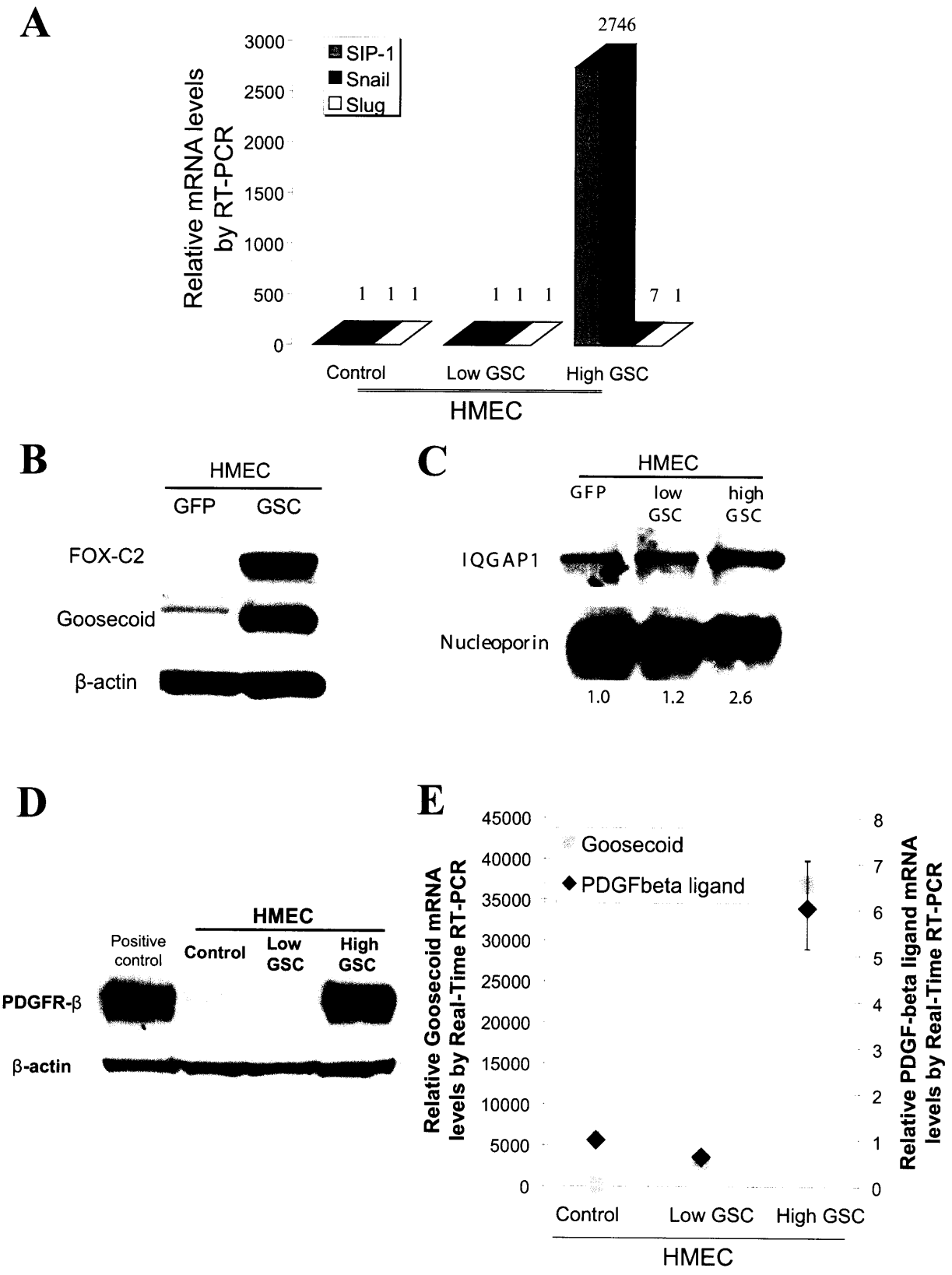

A
Relative mRNA levels by RT-PCR
3000
2500
2000
1500
1000
500
0
SIP-1
Snail
Slug
2746
1 1 1
1 1 1
7 1
Control
Low GSC
High GSC
HMEC
B
HMEC
GFP
GSC
FOX-C2
Goosecoid
β-actin
C
HMEC
GFP
low GSC
high GSC
IQGAP1
Nucleoporin
1.0
1.2
2.6
D
HMEC
Positive control
Control
Low GSC
High GSC
PDGFR-β
β-actin
E
Relative Goosecoid mRNA levels by Real-Time RT-PCR
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
Goosecoid
PDGFbeta ligand
8
7
6
5
4
3
2
1
0
Relative PDGF-beta ligand mRNA levels by Real-Time RT-PCR
Control
Low GSC
High GSC
HMEC

Table 2. Other putative Goosecoid target genes

Shown are a subset of the human genes found to contain putative Gsc binding sites within their promoter. The gene names are accompanied by pertinent information related to this thesis.

tenascin-C ECM component, role in inflammation, wound healing; invasion front expression predicts distant metastases
TGFbeta2 precursor TGFbeta pathway; role in developmental EMT, cancer, wound healing, etc
activin A receptor TGFbeta superfamily
slug (snail homolog 2) role in EMT during development and metastasis
TNFalpha role in EMT and metastasis
protocadherin cadherin superfamily; function and intracellular signal transduction poorly understood
cadherin-like 26
goosecoid literature confirms goosecoid autoregulation
N-cadherin EMT and metastasis
IQGAP1 cytoskeletal rearrangement, upregulated in metastasis screen by Clark et al. 2000
IQGAP2 cytoskeletal rearrangement
angiopoietin-1 antiangiogenic, decreases in cancer
pax6 known to bind goosecoid
FOXF2 regulates several lung-specific genes
FOXD1
MMP16 (matrix metalloproteinase) degradation of extracellular matrix
dkk3 (dickkopf) antagonist of wnt signaling
LISCH7 third most strongly upregulated colonization gene in metastasis screen by Yang et al. 2004

Supplemental Table 1. Characteristics of patients and tumor samples in clinical data set

Available patient information is listed along with the relative *GSC* expression levels in all samples as shown in Figure 1.

ADH, atypical ductal hyperplasia; DCIS, ductal carcinoma *in situ*; IDC, invasive ductal carcinoma; ER, estrogen receptor status; PR, progesterone receptor status; HER2, HER2 protein status; ND, not determined; Pos, positive; Neg, negative

Supplemental Table 1. Characteristics of patients and tumor samples in clinical data set

Case ID	Pathological subtype	Relative GSC level Tumor	maxi error Tumor	mini error Tumor
191	ADH	29.66296	0.37731	0.37257
57	ADH	22.11783	0.39017	0.38341
131	ADH	15.04325	0.55462	0.53490
213	ADH	8.68470	0.61690	0.57599
180	ADH	3.16727	0.12724	0.12233
79	ADH	2.73158	0.04322	0.04255
193	ADH	0.98428	0.05053	0.04807
193	DCIS	59.57957	0.58950	0.58372
191	DCIS	42.32537	3.11005	2.89717
184	DCIS	29.44733	0.43614	0.42978
57	DCIS	28.37041	0.69127	0.67483
170	DCIS	28.09911	0.35851	0.35399
122	DCIS	27.83987	0.88848	0.86100
14	DCIS	26.83985	0.26757	0.26493
198	DCIS	21.58055	0.88717	0.85213
65	DCIS	21.42518	0.88714	0.85187
121	DCIS	19.54444	0.58705	0.56993
11	DCIS	17.72449	0.27409	0.26991
102	DCIS	16.69234	0.44780	0.43610
173	DCIS	13.10603	0.27849	0.27270
133	DCIS	11.39867	0.50241	0.48120
41	DCIS	10.19363	0.76121	0.70832
88	DCIS	9.13030	0.12733	0.12558
60	DCIS	8.60268	0.06385	0.06338
79	DCIS	7.77144	0.38765	0.36923
169	DCIS	5.76677	0.33960	0.32071
5	DCIS	5.73904	0.10685	0.10490
180	DCIS	4.92869	0.04619	0.04576
179	DCIS	4.21639	0.22807	0.21636
131	DCIS	3.48550	0.22402	0.21049
152	DCIS	3.23390	0.12777	0.12291
75	DCIS	3.02933	0.14399	0.13746
112	DCIS	3.02906	0.10034	0.09712
43	DCIS	2.03477	0.05761	0.05602
178	DCIS	1.37681	0.02489	0.02444
12	DCIS	1.01746	0.07617	0.07086
44	DCIS	0.85984	0.00642	0.00638
96	DCIS	0.25088	0.01089	0.01043
89	DCIS	0.17028	0.00903	0.00857
30	DCIS	0.16805	0.00264	0.00260
130	DCIS	0.11315	0.00160	0.00158
183	DCIS	0.10468	0.00376	0.00363
148	DCIS	0.09816	0.00476	0.00454
45	DCIS	0.07914	0.00524	0.00492
72	DCIS	0.01706	0.00152	0.00140
193	IDC	91.26769	1.27881	1.26114

Case ID	Pathological subtype	Relative *GSC* level Tumor	maxi error Tumor	mini error Tumor
170	IDC	34.35150	0.92555	0.90126
14	IDC	29.44764	0.13497	0.13435
198	IDC	16.99501	0.29244	0.28750
102	IDC	16.99021	1.14650	1.07403
121	IDC	13.86517	0.16028	0.15845
122	IDC	11.36775	1.36960	1.22233
41	IDC	7.40581	0.41695	0.39472
65	IDC	6.80730	0.10111	0.09963
88	IDC	5.93324	0.69110	0.61900
112	IDC	5.87579	0.18238	0.17689
43	IDC	5.49973	0.03490	0.03468
180	IDC	4.74298	0.10422	0.10198
173	IDC	3.30778	0.02118	0.02104
169	IDC	2.05162	0.01958	0.01940
79	IDC	1.63739	0.09022	0.08551
133	IDC	0.98251	0.01638	0.01612
178	IDC	0.96344	0.04178	0.04005
44	IDC	0.74527	0.03087	0.02964
30	IDC	0.49488	0.02697	0.02558
96	IDC	0.46472	0.02139	0.02045
5	IDC	0.42156	0.01446	0.01398
131	IDC	0.28468	0.02006	0.01874
179	IDC	0.19471	0.00643	0.00623
130	IDC	0.18013	0.01524	0.01405
153	IDC	0.12232	0.00944	0.00876
148	IDC	0.12169	0.00237	0.00233

Case ID	Relative GSC level Normal	maxi error Normal	mini error Normal
191	2.65492	0.07018	0.06837
57	2.28037	0.04013	0.03943
131	2.02266	0.03464	0.03405
213	1.73455	0.03066	0.03013
180	3.27634	0.05840	0.05738
79	0.26075	0.00729	0.00710
193	1.22778	0.01382	0.01367
193	1.22778	0.01382	0.01367
191	2.65492	0.07018	0.06837
184	0.29037	0.00231	0.00229
57	2.28037	0.04013	0.03943
170	5.93108	0.15917	0.15501
122	0.81602	0.02756	0.02666
14	4.44317	0.14676	0.14207
198	2.10467	0.02835	0.02797
65	0.69470	0.01364	0.01338
121	0.63577	0.03542	0.03355
11	1.90499	0.05829	0.05656
102	2.17029	0.09624	0.09215
173	2.64312	0.08111	0.07869
133	0.63562	0.04581	0.04273
41	0.34343	0.02256	0.02117
88	0.41853	0.01149	0.01118
60	1.19045	0.01648	0.01625
79	0.26075	0.00729	0.00710
169	0.24864	0.00505	0.00495
5	0.50032	0.01372	0.01335
180	3.27634	0.05840	0.05738
179	0.62167	0.00999	0.00983
131	2.02266	0.03464	0.03405
152	2.22917	0.13773	0.12971
75	0.95723	0.02128	0.02082
112	0.42223	0.02060	0.01965
43	0.86221	0.02215	0.02160
178	0.27039	0.01660	0.01564
12	2.01645	0.01822	0.01806
44	0.52332	0.01988	0.01916
96	0.06101	0.00420	0.00393
89	1.67548	0.07813	0.07465
30	0.29072	0.00229	0.00227
130	2.07812	0.03069	0.03025
183	3.02004	0.14900	0.14199
148	1.83760	0.10124	0.09595
45	2.52096	0.05690	0.05565
72	0.30338	0.00513	0.00504
193	1.22778	0.01382	0.01367

Case ID	Relative GSC level Normal	maxi error Normal	mini error Normal
170	5.93108	0.15917	0.15501
14	4.44317	0.14676	0.14207
198	2.10467	0.02835	0.02797
102	2.17029	0.09624	0.09215
121	0.63577	0.03542	0.03355
122	0.81602	0.02756	0.02666
41	0.34343	0.02256	0.02117
65	0.69470	0.01364	0.01338
88	0.41853	0.01149	0.01118
112	0.42223	0.02060	0.01965
43	0.86221	0.02215	0.02160
180	3.27634	0.05840	0.05738
173	2.64312	0.08111	0.07869
169	0.24864	0.00505	0.00495
79	0.26075	0.00729	0.00710
133	0.63562	0.04581	0.04273
178	0.27039	0.01660	0.01564
44	0.52332	0.01988	0.01916
30	0.29072	0.00229	0.00227
96	0.06101	0.00420	0.00393
5	0.50032	0.01372	0.01335
131	2.02266	0.03464	0.03405
179	0.62167	0.00999	0.00983
130	2.07812	0.03069	0.03025
153	1.44793	0.07430	0.07067
148	1.83760	0.10124	0.09595

Case ID	Patient age	Tumor grade	ER	PR	HER2	Nodal status
191	43	N/A	ND	ND	ND	ND
57	36	N/A	Pos	Neg	Neg	Neg
131	37	N/A	Pos	Pos	Pos	Pos
213	45	N/A	ND	ND	ND	Neg
180	46	N/A	Pos	Pos	Neg	Pos
79	54	N/A	Pos	Pos	Neg	Pos
193	45	N/A	Pos	Pos	Neg	Pos
193	45	I	Pos	Pos	Neg	Pos
191	43	II	ND	ND	ND	ND
184	54	I	ND	ND	ND	ND
57	36	I	Pos	Neg	Neg	Neg
170	44	II	Pos	Pos	Pos	Pos
122	45	II	Pos	Pos	Neg	ND
14	44	I	Pos	Pos	ND	Pos
198	30	II	Pos	Pos	Neg	Neg
65	39	III	Pos	Pos	Neg	Neg
121	45	II	Pos	Pos	Pos	Pos
11	49	I	ND	ND	ND	ND
102	55	I	Pos	Neg	Neg	Pos
173	52	I	Pos	Pos	Neg	Neg
133	44	III	Neg	Neg	Pos	Pos
41	55	II	Pos	Pos	ND	Neg
88	35	III	Pos	Pos	ND	Pos
60	48	II	ND	ND	ND	ND
79	54	I	Pos	Pos	Neg	Pos
169	34	II	Pos	Pos	Neg	Pos
5	48	II	Pos	Pos	ND	ND
180	46	I	Pos	Pos	Neg	Pos
179	37	III	Neg	Neg	Pos	Pos
131	37	II	Pos	Pos	Pos	Pos
152	55	III	ND	ND	ND	Neg
75	63	II	ND	ND	ND	ND
112	31	III	Neg	Pos	Neg	Pos
43	53	II	Pos	Neg	Neg	Pos
178	43	III	Pos	Pos	Pos	ND
12	39	I	ND	ND	ND	ND
44	28	III	Pos	Pos	Neg	Neg
96	31	III	Neg	Neg	Neg	Pos
89	35	III	ND	ND	ND	ND
30	47	III	Neg	Neg	Neg	Pos
130	54	II	Pos	Pos	Neg	Pos
183	46	II	ND	ND	ND	Pos
148	42	II	Pos	Pos	Neg	Pos
45	36	I	Pos	Neg	Neg	Neg
72	42	II	ND	ND	ND	ND
193	45	I	Pos	Pos	Neg	Pos

Case ID	Patient age	Tumor grade	ER	PR	HER2	Nodal status
170	44	II	Pos	Pos	Pos	Pos
14	44	I	Pos	Pos	ND	Pos
198	30	II	Pos	Pos	Neg	Neg
102	55	I	Pos	Neg	Neg	Pos
121	45	II	Pos	Pos	Pos	Pos
122	45	II	Pos	Pos	Neg	ND
41	55	II	Pos	Pos	ND	Neg
65	39	III	Pos	Pos	Neg	Neg
88	35	III	Pos	Pos	ND	Pos
112	31	III	Neg	Pos	Neg	Pos
43	53	II	Pos	Neg	Neg	Pos
180	46	I	Pos	Pos	Neg	ND
173	52	I	Pos	Pos	Neg	ND
169	34	II	Pos	Pos	Neg	Pos
79	54	I	Pos	Pos	Neg	Pos
133	44	III	Neg	Neg	Pos	Pos
178	43	III	Pos	Pos	Pos	Pos
44	28	III	Pos	Pos	Neg	Neg
30	47	III	Neg	Neg	Neg	Pos
96	31	III	Neg	Neg	Neg	Pos
5	48	II	Pos	Pos	ND	ND
131	37	II	Pos	Pos	Pos	Pos
179	37	III	Neg	Neg	Pos	Pos
130	54	II	Pos	Pos	Neg	Pos
153	46	I	Pos	Pos	Pos	Pos
148	42	II	Pos	Pos	Neg	Pos

Materials and Methods

RNA preparation and RT-PCR

The clinical cohort examined was previously described (Ma et al. 2003). 72 tumor samples were obtained from 40 patients, 28 of whom had two or more pathological subtypes of breast cancer detectable at diagnosis, and each was accompanied by a patient-matched normal breast tissue sample. The Massachusetts Institute of Technology Committee on the Use of Humans as Experimental Subjects and the Massachusetts General Hospital Human Research Committee approved this study of deidentified samples. cDNAs from the previous study were additionally analyzed for *GSC* by real-time quantitative PCR analysis using the ABI 7900HT system as previously described (Ma et al. 2003). The sequences of the *GSC*-specific fluorogenic MGB probe (5' to 3') and the PCR primer pair, respectively, were as follows: VIC-CCCACCGTAGTATTTAT, GCCGCCCGCGACTAG, and CACTTTATTGTACTGTCACCCTTAATTTAAC. Statistical significance was calculated for this clinical data set using the paired Student's *t*-test, and relative expression was calculated as described (Yang et al. 2004).

For cell line analyses, total RNA was purified using RNA STAT-60 (Tel-Test) and RNase-free DNase set (Qiagen) according to manufacturer's instructions. Hexanucleotide mix (Roche) was used for reverse transcription. Quantitative real-time RT-PCR was performed in triplicate using the iCycler apparatus (Bio-Rad) and SYBR-Green detection reagent, either from stock (Molecular Probes) or in commercial master mix (Perkin-Elmer Applied Biosystems). The sequences of the *GSC*-specific primer pairs were (5' to 3') TCTCAACCAGCTGCACTGTC (left) and GGCGGTTCTTAAACCAGACC (right), and that of the *GAPDH*-specific pairs were AGCCACATCGCTCAGACAC (left) and AATGAAGGGGTCATTGATGG (right).

Experimental data was normalized to *GAPDH* and relative expression calculated as described (Yang et al. 2004).

Expression constructs and virus generation

Full-length mouse *Goosecoid* cDNA (Danilov et al. 1998) provided by Dr. Martin Blum (Forschungszentrum Karlsruhe, Karlsruhe, Germany) was subcloned with or without an HA-antigen tag at the amino terminus into the pWZL-Blasticidin vector. A corresponding vector containing the GFP gene was used as control. ΔN90 β-catenin consisting of mouse β-catenin containing amino-terminal deletions of 90aa (Barth et al. 1997) and Lef-vp16 consisting of mouse Lef-1 fused to the transactivation domain from the Herpes Simplex Virus VP16 protein (Aoki et al. 1999) provided by Dr. Masahiro Aoki (Scripps Research Institute, La Jolla, CA) were expressed using the pBabe-Puromycin vector. Activated, myc-tagged, human TGFβ type I receptor cDNA (Wieser et al. 1995) provided by Dr. Joan Massagué (Sloan-Kettering Institute, New York, NY) was expressed using the pWZL-Blasticidin vector. pWZL and pBabe amphotropic viruses and lentiviruses were generated and used for target cell infection as previously described (Stewart et al. 2003). Lentiviral shRNA-expressing constructs were designed and used as described against (Stewart et al. 2003). The construct that gave significant knockdown was directed against the sequence ACTTACCTAACTCGAAGGACT within the human *Goosecoid* gene.

Cell culture

The Madin-Darby canine kidney (MDCK) cell line was obtained from ATCC and cultured in Dulbecco's modified Eagle medium supplemented with 10% heat-inactivated fetal calf serum.

The immortalized, non-transformed human mammary epithelial cell (HMEC) line, expressing the SV40 early region and *hTERT,* was previously described (Elenbaas et al. 2001) and cultured in Dulbecco's modified Eagle medium and F12 medium (1:1) containing the supplements EGF (10 ng/ml), insulin (10 μg/ml), and hydrocortisone (0.5 μg/ml), with noted exceptions. The Gsc-expressing HMEC cells were generated using differential trypsinization of the polyclonal population of Gsc-transduced cells to separate out the scattered, less adherent cells from those not expressing substantial amounts of Gsc, as confirmed by western blotting and quantitative RT-PCR. Soluble, activated TGF-β1 ligand (R&D Systems) was used at a working concentration of 100pM, or 2.5ng/ml, in the presence of 5% calf serum. The MDA-MB-231 cell line was maintained in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum.

Antibodies, immunoblotting, and immunofluorescence

A rabbit polyclonal antibody against Gsc was generated using a KLH-conjugated peptide of the sequence CSENAEKWNKTSSSKA, and resulting antisera were affinity purified (Covance). The specificity of the antibody was confirmed by western immunoblotting using whole cell lysates expressing either tagged or untagged ectopic Gsc. Other primary antibodies used were fibronectin (BD Transduction Labs and 297.1 antibody courtesy of Sophie Snitkovsky and Richard Hynes, MIT Center for Cancer Research, Cambridge, MA), vimentin (V9 Neomarkers #MS129P), N-cadherin (Zymed #180224, BD Transduction Labs #610920), PDGFR-β (Research Diagnostics #pdgfrbabrx), β-actin (Abcam #8226), pan-cytokeratin (Biogenex #071M), α-catenin, γ-catenin, β-catenin and E-cadherin (BD Transduction Labs #C21620, #610254, #C19220, #610182). Immunoblotting for IQGAP1 was performed as described (Frew 2004). Standard procedures were used for immunoblotting and immunofluorescence.

Transwell migration assays

Cells were plated on cell culture inserts (Falcon) containing a filter with 8.0 micrometer pores. Total cells and migrated cells were quantified using crystal violet staining after time indicated and compared to control for differences in cell number as described (Clark et al. 2000).

Mice and injection of tumor cells

Female NOD-SCID mice (propagated on site), and nude mice (NCR nude, Taconic) were used in these studies and all protocols were approved by the Massachusetts Institute of Technology Committee on Animal Care. Nude mice received 400 rad of γ-radiation using a dual ^{137}Cesium source one day prior to tumor cell injection. Mice were anesthetized with either avertin (IP) or with isoflurane (inhalation). For orthotopic injections, one million cells in 30μL of Matrigel (Becton-Dickinson) diluted 1:2 in medium were injected into each of two mammary glands per NOD-SCID mouse. For subcutaneous injections, 2x10^6 cells in 160μL of Matrigel diluted 1:2 in medium were injected at each of three sites per nude mouse. For tail vein injections, 2x10^6 cells in 200μL PBS were injected per mouse. Tumor diameters were measured multiple times per week using precision calipers.

Visualization and quantification of GFP-labeled lung metastases

Upon necropsy, lungs of injected mice were removed, separated into individual lobes, and examined under a Leica MZ 12 fluorescence dissection microscope. Images of both faces of all lobes were captured at identical settings, and the fluorescent metastatic nodules in each image were analyzed using CellProfiler image analysis software developed in the laboratory of Dr.

David Sabatini (www.cellprofiler.org(Carpenter et al. 2006)). The unpaired Student's *t*-test was used for statistical comparisons of these data.

Acknowledgments

I thank I. Ben-Porath, J. Yang, S. Mani, C. Kuperwasser, B. Elenbaas, R. Hynes, J. Lees, T. Ince, S. McCallister, R. Lee , A. Orimo, L. Spirio, C. Scheel, A. Karnoub, S. Stewart, and other members of the Weinberg lab for invaluable input during the course of this work. I also thank The National Cancer Institute, M. Blum, R. Hynes, J. Yang, S. Mani, M. Aoki, J. Massagué, M. van de Wetering, B. Elenbaas, C. Kuperwasser, and L. Spirio for reagents, M. Brooks, Tom DiCesare, Sanjeev Pillaik, S. Astrof, L. Xu and M. Rockas for technical assistance, and G. Bell for help with statistical analysis. I acknowledge the support of the W.M. Keck Biological Imaging Facility at Whitehead Institute. K.A.H. is a recipient of a US Army Pre-doctoral Breast Cancer Fellowship. R.A.W. is a Daniel K. Ludwig Cancer Research Professor and an American Cancer Society Research professor. This research was supported by NIH grant R01-CA078461 and by a grant from the Breast Cancer Research Foundation. The authors declare that they have no conflicting financial interests.

Chapter 3

Conclusions and Future Directions

A novel role for the developmental gene *Goosecoid*

In summary, this work has identified a role for the Goosecoid homeobox transcription factor in cancer, specifically in promoting tumor cell malignancy. *GSC* mRNA was found to be expressed in a majority of microdissected tumor tissues at levels significantly above the levels found in patient-matched, microdissected normal control tissue. To facilitate subsequent studies of this developmental gene in disease, a rabbit polyclonal antibody was generated with specificity to Gsc. Ectopic expression of Gsc in human breast and canine kidney cells was found to generate a shift in cellular morphology and in the composition of the cytoskeleton and cell-cell adhesion complexes. This shift was accompanied by increased cell motility, and indicates that Gsc induces an epithelial-mesenchymal transition (EMT), a transdifferentiation event that is known to promote tumor invasion and metastasis.

I observed that Gsc reduces the growth rate of most cells into which it is introduced. While *GSC* mRNA expression is significantly upregulated in clinical breast tumors, its expression was surprisingly rare in an extensive panel of human cancer cell lines and did not fit a clear pattern. The expression of *Gsc* mRNA was detectable, however, in a set of increasingly metastatic melanoma cell lines and correlated with enhanced aggressiveness.

GSC was found to be inducible by TGF-β signaling in human mammary epithelial cells (HMECs), as has been shown in *Xenopus* embryos, and its expression correlated with the induction of a morphology change and an EMT. *GSC* was not found to be inducible by Wnt/β-catenin signaling in these cells, surprisingly. Other putative signals that may potentially regulate *GSC* per my analysis of its promoter sequence include certain growth factors and the TNFα/NFκB pathway.

Analyses into the signaling downstream of Gsc revealed that not only are the classical hallmarks of an EMT induced in Gsc-expressing HMECs but other relevant cancer genes as well. Gsc robustly upregulates the E-cadherin silencing transcription factor *SIP1* in HMECs but not *SNAI1(Snail)* or *SNAI2 (Slug)*. Other genes found to be induced by Gsc that likely have roles in cell plasticity and metastasis include *Foxc2*, another developmental transcription factor that promotes the formation of mesenchymal tissues, and IQGAP1, a known regulator of the cytoskeleton, cell-cell adhesion and cell polarity. Finally, the PDGFR-β and the corresponding *PDGF-β* ligand were also induced by Gsc.

Moreover, Gsc significantly enhanced the ability of breast cancer cells to form pulmonary metastases in mice as assessed by the experimental metastasis assay (tail vein injection), without increasing the rate of tumor growth at the subcutaneous and orthotopic sites. *In vitro*, Gsc expression pushed the morphology of these cells to a more mesenchymal state and enhanced cell motility, suggesting that a Gsc-induced transdifferentiation contributed to the enhanced metastatic efficiency observed *in vivo*.

In conclusion, my results demonstrate that the Goosecoid homeobox transcription factor, a major orchestrator of Spemann organizer biology during gastrulation, plays an important role in activating cell properties associated with tumor progression to malignancy. Moreover, given the role of the EMT in breast development (specifically in mammary gland branching morphogenesis), wound healing, and fibrosis, Goosecoid may have an role in these processes as well.

Additional embryonic transcription factors promote the EMT and metastasis

During the course of my described studies, other developmental transcription factors, including SNAI1 (Snail), SNAI2 (Slug) and Twist, have also been linked to the EMT and tumor invasion as previously mentioned (Savagner et al. 1997; Batlle et al. 2000; Cano et al. 2000; Yang et al. 2004). While the concept of the EMT as a driving force behind human cancer metastasis is well described, there are still very limited *in vivo* data demonstrating that such genes inducing the mesenchymal state contribute functionally to tumor metastasis (Thiery and Sleeman 2006). Suppression of Twist expression *in vivo*, in a xenograft model of cancer progression previously described, was found to antagonize tumor cells' ability to generate metastases from a primary tumor(Yang et al. 2004). Snail was found to be a potent driver of tumor recurrence in a conditional transgenic mouse model expressing activated ERBB2. Interestingly, the recurrent tumors consisted of cells with a mesenchymal morphology (Moody et al. 2005). These data and my own *in vivo* studies of Goosecoid strongly support a role for EMT-inducing embryonic transcription factors in tumor metastasis.

Where within the metastatic cascade does Goosecoid function?

Here, I have found that Gsc is sufficient to enhance metastatic behavior in an *in vivo* model of experimental metastasis. Gsc may augment metastatic colonization by promoting extravasation, cell survival in the environment of the lung, or migration to hospitable microenvironments within the lung. Ideally, I would like to have identified and employed a cell line in which Gsc augmented metastasis from a primary tumor site, so that I could comprehensively assess which steps of metastasis Gsc promotes. The described preliminary experiments using the A375 human melanoma cell line may have identified such a cell line.

However, even if I were unable to identify an existing cancer cell line that can become more metastatic at the primary tumor site upon *Gsc* expression, I would not be able to rule out the possibility that Gsc does indeed promote even very early steps in the metastatic cascade. The EMT is thought to have a role in the early as well as the late stages of metastasis. The initial detachment of tumor cells from each other and from their ECM anchors and their invasion through the basement membrane likely requires at least a partial transdifferentiation towards a mesenchymal phenotype, for example (Thiery 2002). The EMT state may also be permissive to active intravasation and to anoikis resistance. Snail has already been shown to protect cells from this type of apoptosis (Vega et al. 2004). It may simply be difficult to find a cell line with the genetic background and *in vivo* behavior necessary to reveal such a function for Gsc.

The experimental metastasis assay (tail vein injection) allows only for the assessment of the later steps of metastasis- survival in circulation, arrest in the lung capillaries, extravasation into the lung parenchyma, proliferation in the foreign lung environment, and evasion of host immune surveillance. Because Gsc-expressing cells augment metastasis by this assay, Gsc likely promotes one or more of these steps.

My injection of the Gsc-expressing and control MDA-MB-231 cells into the subcutaneous and orthotopic sites of mice did not shed much additional light on these questions. My motivation for monitoring the effects of Gsc on the growth of primary tumors was to address the possibility that Gsc appears to foster metastasis upon tail vein injection merely by universally increasing cell proliferation in any *in vivo* setting. In fact, I instead observed that Gsc reduced the proliferation rate of cancer cells *in vivo*, making it highly unlikely that its mechanism of enhancing metastasis derives from an ability to favor the proliferation of cells in sites of dissemination. Moreover, as previously mentioned, the quantification of the size of the lung

metastases generated following tail vein injection revealed that the metastases formed by Gsc-expressing cells were not significantly larger than control metastases.

Because the expression of Gsc decreased the size of primary tumors, I could not accurately compare the effects of Gsc on the early steps of metastasis from this site. Gsc-expressing cells were likely being selected against within the primary tumors. Moreover, the smaller Gsc tumors would be expected to shed fewer tumor cells into the circulation simply as a result of this size difference, and any enhancement of metastasis by Gsc would therefore have to overcome these differences in order to be observable. Not unsurprisingly, therefore, when I examined the prevalence of lung metastases that eventually arose in these mice, an enhancement of metastasis in the mice with the smaller Gsc tumors was not observed.

Interestingly, experiments performed by Minn and colleagues identified two distinct classes of genes that promote breast cancer metastasis to the lungs (Minn et al. 2005). Both classes of genes promote metastatic colonization as assessed by the experimental tail vein assay. However, one class not only confers metastatic capabilities but also functions within the primary tumor to augment breast tumorigenicity. Cells expressing such genes promote primary tumor growth when injected into the orthotopic site. The second class, in contrast, specifically confers metastatic capabilities but no growth advantage at the primary tumor site. Cells expressing such genes were found to be rare within the original tumor cell population and are thought to act mainly as virulence genes. These rare cells are likely then selected for when they reach the lung environment (Minn et al. 2005).

Accordingly, since Gsc increases the ease with which MDA-MB-231 breast cancer cells form lung metastases but does not enhance primary tumor formation, this gene likely falls within this second class of metastasis-promoting genes. Perhaps, in certain types of cancer, *GSC* is

upregulated as a result of contextual signals in only the rare cells receiving relevant signals from surrounding stromal cells, as proposed earlier. In other types of cancer, such as the ductal-type carcinomas as examined here, *GSC* expression may instead be quite widespread in the primary tumor cell population. Perhaps in these cases, *GSC* is induced merely as a side effect of particular signaling that primarily confers a selective advantage during tumorigenesis. In time, this side-effect can acquire functional relevance, conferring a selective advantage for metastasis.

I do not yet know whether Gsc is essential for metastasis. The effect of *GSC* knockdown in aggressive tumor cells (e.g. cell lines Sum 1315, MC-1 and MC-2) on the occurrence of metastasis *in vivo* remains unclear, as these experiments have proven to be technically challenging. It should be noted, however, that the inhibition of *GSC* may not revert cells back to a less aggressive epithelial state and that such an observation would be an uninformative negative result. Gsc might initiate the EMT but not sustain it. Other types of signaling, such as TGF-β autocrine signaling, might stabilize the EMT state thereafter, rendering Gsc no longer necessary (Oft et al. 1998). In fact, it is known that Gsc can bind its own promoter, silencing further expression by way of a negative feedback loop (Danilov et al. 1998).

Early changes in gene expression may drive clinical metastasis

Our findings that *GSC* expression is upregulated in the vast majority of clinical ductal-type tumors supports a role for this embryonic transcription factor in human breast cancer. The upregulation of *GSC* occurs quite early in multi-step cancer progression rather than concurrently with the overt display of the invasive phenotype. This result is not unusual for breast carcinoma progression; for example, the *HER2/neu* gene, known to promote invasive cell behavior (Holbro et al. 2003) and routinely used to inform both patient treatment and prognosis, is likewise already

overexpressed in human tumors prior to the overt onset of invasiveness (Menard et al. 2001). Our observations are in accord with other gene expression studies examining different stages of ductal-type breast cancer progression. These studies have shown that most expression changes associated with invasiveness are already present in pre-invasive tissue (Ma et al. 2003; Porter et al. 2003).

Moreover, our observations are consistent with other published results demonstrating that several genes shown to promote the metastatic behaviors and poor prognosis of aggressive cancers, such as *Slug* and *HOXB13*, are expressed in clinical specimens prior to the appearance of the malignant tumor phenotype (Ma et al. 2004; Gupta et al. 2005a). Thus, it possible that in human ductal-type breast tumors, Gsc primes cells for the expression of aggressive phenotypes, which manifest themselves only later, in the context of subsequent alterations.

Importantly, our observations relate to the question of where within the course of multistep tumorigenesis tumor cells acquire metastatic proclivity. Our observations of the phenotypic and functional effects of Gsc expression in epithelial cells strongly argue that Gsc can serve as a functional contributor to the metastatic phenotype in clinical tumors. *GSC* upregulation above a certain threshold level may denote which early-stage tumors are primed for metastasis. In other words, it may denote which tumors are capable of becoming metastatic if they eventually incur the appropriate additional cooperative lesions necessary for the clinical manifestation of this phenotype. Gsc may therefore be essential to the metastatic phenotype of some tumors but not sufficient.

Unlike most metastasis genes studied in the context of clinical tumors, the degree to which *GSC* is upregulated above normal levels in ductal-type tumors does not incrementally increase from the least to the most aggressive breast tumor subtypes. One possible explanation is

that levels of *GSC* above a certain threshold value may denote tumor cells that are in a certain cell state or have a particular genetic background, and that as long as certain additional mutations occur in the context of such a background, metastasis can result. In other words, while an EMT and other metastatic changes are not overtly manifested early in tumorigenesis, *GSC* expression may distinguish those cells that are poised in a cell state that is fertile for metastasis from those cells that are not.

In summary, our observations support a model in which metastasis results from the clonal selection of cells harboring certain late-stage genetic lesions which function cooperatively with the genetics present in primary tumors from very early stages of tumorigenesis.

It is important to note that the quantification of *Goosecoid* at the RNA level in these clinical samples gives an incomplete assessment of Goosecoid activity. This transcription factor may be regulated at the protein level by post-translational modification(s), for example. Interestingly, Snail, another developmental transcription factor implicated in the EMT and metastasis as previously mentioned, is known to be regulated by GSK-3β, as is β-catenin. GSK-3β binds and phosphorylates Snail at two different motifs, thereby regulating both its stability and subcellular localization (Zhou et al. 2004). A similar mechanism may exist for Goosecoid that dictates, at least in part, its functional activity. It may be, therefore, that an examination of the status of Goosecoid at the protein level in clinical samples would reveal a significant correlation between Goosecoid activity and the invasive phenotype.

EMT-inducing genes as proliferation suppressors

In my hands, Gsc reduced the growth of most of the cell lines I ectopically expressed it in, both *in vitro* and *in vivo*. This was routinely observed in non-transformed cell lines. Rather

than being an overexpression artifact, this effect may have notable significance. Recent analyses in the literature suggest an inherent incompatibility with cell proliferation and migration. In one study, Wang and colleagues compared the gene expression signature of invasive cells collected using an *in vivo* chemotactic invasion assay to that of cells within the general primary tumor cell population. Three classes of genes in particular were found to be differentially expressed between these two groups of cells, revealing that actively invading tumor cells *in vivo* are generally (1) less proliferative, (2) more resistant to apoptosis, and (3) more actively reconfiguring their cytoskeleton compared to the bulk of tumor cells within a primary tumor (Wang et al. 2004). In another study, Snail was found to decrease the cell cycle by directly repressing *Cyclin D2* transcription. Snail also conferred resistance to apoptosis when ectopically expressed in cultured epithelial cells and in developing chick embryos (Vega et al. 2004).

Together, these observations raise a conundrum- invasive and metastatic tumor cells may in fact be less proliferative than their more stationary tumor cell neighbors. How, then, are they maintained within a primary tumor cell population that is continuously evolving through successive rounds of clonal selection and expansion? One possibility is that even though genes promoting invasion may decrease the proliferation rate of cells expressing them, they may confer enough of an advantage in other ways that they are selected for nevertheless (Arendt et al. 2001). Resistance to apoptosis may be one such selective advantage.

Alternatively, or in certain cases, invasive and metastatic cells may constitute only a subfraction of cells within a tumor, possibly the cells in close juxtaposition with stromal cells, for example, which transiently enter the EMT state. Such cells may revert back to their original epithelial phenotype upon reaching distant organs, where the original contextual signals at the primary tumor site are presumably now absent. Indeed, the EMT state is known to be reversible

(Grunert et al. 2003). Mesenchymal tumor cells in distant organs may revert back to their epithelial state, as has been observed during development, through a process known as mesenchymal-epithelial transition (MET) (Hay and Zuk 1995). An MET may allow invasive and metastatic cells to thus return to a more proliferative state, which may be important for colonization. It remains to be determined whether the EMT state induced by Gsc in the human and canine epithelial cells is reversible. An inducible Gsc construct would be especially useful in answering this and other questions.

Thus, even though Gsc may in fact limit the proliferation rate of cells, its elevated expression may still be selected for in certain types of tumors. Perhaps the lung metastases generated in my experimental metastasis (tail vein) assay would actually have grown much faster were Gsc only transiently rather than stably expressed. Mechanistically, Gsc might alter the proliferation rate and apoptotic sensitivity of HMECs by inducing *Snail*, which it was observed to induce only mildly, or by inducing *SIP1*. Gsc induced *SIP1* far more robustly, but *SIP1* has not yet been linked to apoptosis resistance. A knockdown Snail construct would be necessary to determine whether the induction of *Snail* by Gsc is functionally significant.

Goosecoid as a possible cancer stem cell marker

Gsc may promote a cancer stem cell state in cells that express it, as may all EMT-inducing genes. Cancer stem cells are defined as the subset of cells within a primary tumor that are able to drive tumor formation and expansion (Reya et al. 2001). Solid tumors have been shown to consist of a heterogeneous population of cell variants which are not all equally capable of giving rise to new tumors (Al-Hajj et al. 2003). Those variants found within a heterogeneous tumor that are capable of generating new tumors are deemed the tumor stem cells. Tumor stem

cells are so named because they are thought to possess the characteristics that are attributed to normal stem cells, including the power of self renewal. Indeed, tumors are generally clonal in nature yet they persistently consist of both non-tumorigenic and tumorigenic subpopulations, suggesting they do harbor a self-renewing cell subpopulation that can generate self as well as differentiated variants (Reya et al. 2001).

Two recent review articles have proposed the possibility that the EMT state identifies the subpopulation of stem cells within a tumor, but solid experimental evidence supporting this hypothesis is still lacking (Brabletz et al. 2005; Prindull 2005). Overlap does exist between the types of signals known to be involved in inducing an EMT and in maintaining stemness, and examples include Wnt, Notch and Hedgehog signals (Brabletz et al. 2005; Huber et al. 2005). Moreover, if cells that have undergone an EMT are to seed metastases, as has been hypothesized, such cells would presumably need to possess stem cell characteristics including unlimited replication potential (Thiery 2002; Brabletz et al. 2005).

The prevalence of *GSC* overexpression in the ductal-type tumors suggests, among others, the possibility that *GSC* may serve as a marker of a cell subpopulation especially capable of generating and sustaining tumors upon transformation. Such cells may initially have been quite rare within normal tissue, consistent with the minimal expression level of *GSC* in the normal samples analyzed. However, as a tumor cell arose and proliferated, so too did this specific subpopulation of cells that was originally quite rare, such that now *GSC* effectively appears to be overexpressed across the tumor cell population relative to normal tissue (Figure 1B). In other words, it remains unclear whether *GSC* serves as a passive marker of a target population that is particularly susceptible to transformation. Such a target population might in fact be the normal cell type from which tumor stem cells can arise. Interestingly, the embryonic cell population

expressing Gsc at the anterior of the primitive streak in the mouse is thought to have stem cell characteristics, and Gsc might functionally contribute to such cellular traits (Blum et al. 1994).

Specific effects of *Goosecoid* overexpression in clinical ductal-type tumors

If Gsc were not merely a marker but play an active, functional role in the pathogenesis of ductal-type tumors, its expression above a certain threshold may in fact promote metastasis as hypothesized. Given the described phenotypes observed in human and canine epithelial cells, I would expect Gsc to be promoting an EMT. Perplexingly, however, ductal-type tumors are not generally seen to lose epithelial E-cadherin expression and do not show a morphological shift suggestive of such a transdifferentiation (Moll et al. 1993).

Nevertheless, we cannot rule out the possibility that Gsc does, at least in part, function in ductal-type carcinomas to destabilize the epithelial cell phenotype. This is because we cannot rule out the possibility that a subtle, partial EMT occurs. Careful examination of ductal-type tumors revealed that 54% of poorly differentiated, more highly malignant IDCs showed reduced E-cadherin staining, manifested as heterogeneous staining dotted over the cell borders (Moll et al. 1993). Another study confirmed that while present in all *in situ* ductal and invasive ductal carcinomas examined, E-cadherin staining had a patchy distribution of variable intensity (Rasbridge et al. 1993). These observations demonstrate that a subtle reduction in E-cadherin protein does occur within ductal-type tumors. We do not yet know whether this subtle reduction has any functional role in the metastasis of ductal-type tumors.

The other biological functions of Gsc may also be important, such as a potential role for this protein in apoptosis resistance and its role in enhancing motility. Notably, the observed enhancement of motility may not be a direct effect of Gsc expression and may be contingent

upon an EMT. If Gsc can directly enhance motility, then its expression in ductal-type tumors above a certain threshold level may promote alternative types of cell movement, such as cohort migration. Cohort migration is a type of tumor cell dispersal in which tumor cells detach from a primary tumor in groups rather than as individual mesenchymal cells (Christiansen and Rajasekaran 2006).

Unfortunately, the clinical data set used in our studies of *GSC* expression does not include patient outcome information. While *GSC* expression levels did not correlate with lymph node positivity in these samples, we cannot know whether or not there was a relationship between *GSC* expression and the eventual occurrence of distant metastases. In breast cancer, however, lymph node status is an established prognostic factor (Weigand et al. 1982). It may be additional mutations that are incurred, functioning in cooperation with *GSC* upregulation, that more accurately predict patient outcome.

Goosecoid expression patterns in other tumor types

I would very much like to know what the pattern of *GSC* expression looks like in a variety of other cancer types. While the biological mechanisms by which different types of carcinomas develop are quite similar, significant differences do exist between the biologies of carcinomas derived from different types of epithelia. In tumors derived from epithelia other than the breast or that belong to an alternative class of breast cancer, such as lobular breast cancer, *GSC* might prove to unambiguously correlate with invasiveness, E-cadherin loss, and/or clinical prognosis. Certain subtypes of gastric cancer, for example, have been found to overexpress other prominent EMT-inducing embryonic transcription factors, including SIP1, Snail, and Twist (Rosivatz et al. 2002). *SIP1* levels were found to closely correlate with E-cadherin levels in

intestinal type gastric cancer samples, but *Twist* and *Snail* were not upregulated in this tumor subtype. N-cadherin and Twist levels were closely correlated in a subset of diffuse type gastric cancer samples, and *Snail* also correlated with E-cadherin loss in diffuse type samples (Rosivatz et al. 2002). It would be especially interesting to know whether *GSC* is upregulated in the samples that have reduced E-cadherin expression but no upregulation of SIP1, Snail or Twist.

For reasons previously discussed, it would also be interesting to assess whether *GSC* is upregulated in the population of tumor cells at the invasive edge of colon tumors found by Brabletz and colleagues to show evidence of an EMT and enhanced Wnt signaling (Brabletz et al. 2001). An antibody (or an optimized protocol using our antibody) that could be used to specifically detect Gsc expression *in situ*, in clinical tissues, would be invaluable for these studies.

Crosstalk and redundancy among the EMT-inducing embryonic transcription factors

As detailed earlier, since my thesis work on Gsc began, seven additional embryonic transcription factors were found capable of inducing an EMT and/or metastasis. These factors include Snail, Slug, SIP1(ZEB2), Foxc2 (MFH-1), Twist, δEF1 (ZEB1), and E47 (E12) (Savagner et al. 1997{Gupta, 2005 #67; Batlle et al. 2000; Cano et al. 2000; Comijn et al. 2001)Yang, 2004 #62}(Perez-Moreno et al. 2001; Eger et al. 2005). As discussed, all except Foxc2 are known to repress *E-cadherin* expression by binding specific E-boxes of the proximal promoter (Hajra et al. 2002; Bolos et al. 2003) (Eger et al. 2005) (Perez-Moreno et al. 2001; Yang et al. 2004). It remains to be determined whether Gsc likewise binds to these E-boxes in the *E-cadherin* promoter, but the absence of a putative Gsc binding site in this region of the genome suggests that it does not. Alternatively, Gsc could destabilize the epithelial phenotype by

inducing other signals, such as IQGAP1, a TGF-β autocrine loop, TNFα signaling, and/or MMPs, all of which are potentially direct Gsc targets by my analysis (Figure 2B).

One provocative question is whether these various transcription factors are functionally equivalent, making their coexpression in a given setting redundant. These factors may work cooperatively, and they may also regulate each other's expression. Gsc did strongly induce *SIP1* and FOXC2 in HMECs, and *Snail* only very mildly, but not Twist or *Slug*. I do not know whether these genes are direct transcriptional targets of Gsc or merely reflective of the EMT state that ectopic Gsc induces in HMECs. Unpublished work completed in collaboration with Jing Yang and Sendurai Mani in our lab demonstrates that Twist does not induce the same subset of these transcription factors in HMECs as Gsc (data not shown), and that Gsc and Twist do not induce each other in these cells (data not shown). Moreover, we have observed that FOXC2 similarly does not induce the same spectrum of EMT markers that Gsc and Twist do (data not shown). Together, these observations indicate that these various EMT-inducing embryonic transcription factors can cross-communicate to some degree and that there is incomplete overlap in the sets of genes they induce.

We have also observed that the expression patterns of *GSC*, *Twist* and *FOXC2* in various cancer cell lines are distinct, though not without some overlap. Surprisingly, the expression of *GSC* was found to be quite rare in established cancer cell lines, in contrast to the seven other factors named (Batlle et al. 2000; Cano et al. 2000; Comijn et al. 2001; Bolos et al. 2003; Eger et al. 2005) (and data not shown). *GSC* expression in such cell lines does not correlate with E-cadherin or invasive status, also unlike the other transcription factors named. I do not yet understand why this is the case.

Another key next step in this work would be to broadly examine the expression patterns of these various transcription factors in clinical tumor samples of many types. Also, it would be especially interesting to assess whether one or more of the other transcription factors is expressed in the ~30% of the ductal-type tumors examined here that did not contain elevated *GSC* mRNA. Finally, I would be very interested in knowing whether these various transcription factors are regulated by distinct or common signals.

Signaling upstream of *Goosecoid*

Other pressing questions remain unanswered. What induces *GSC* upregulation in a clinical setting? I have shown that *GSC* can be induced by TGF-β signaling in human breast epithelial cells, but do not yet know whether this is the means by which *GSC* expression is elevated in clinical tissues. Is *GSC* only transiently induced in a subset of cells within a primary tumor as a result of exposure to stromal signals at the tumor edge? Such cancer cells would likely no longer express *GSC* upon migrating to distant tissues, having presumably abandoned their stromal cells neighbors and thus the source of the contextual signals. Or is *GSC* stably upregulated instead? In other words, does *GSC* upregulation result from contextual signals, the stable misexpression of genes upstream to *GSC*, and/or changes to the *GSC* gene sequence at the DNA level (genetic or epigenetic)? In my favorite hypothetical model, *GSC* is induced at the invasion front of certain tumors as a result of contextual signals (Figure 1A). While I do not yet know the validity of this specific hypothesis, I have demonstrated here that Gsc can promote tumor metastasis and therefore warrants further study by cancer biologists.

The degree to which *GSC* is induced in HMECs by TGF-β signaling is fairly low, suggesting that if TGF-β signaling upregulates *GSC* expression in tumors, it likely does so in

cooperation with additional signaling from another pathway, as is the case in development (Watabe et al. 1995). In *Xenopus*, Wnt/β-catenin signaling first induces Siamois or Twin, proteins that in turn sit on the Wnt-responsive element in the *GSC* promoter (Watabe et al. 1995; Nishita et al. 2000). It was a surprise that *GSC* was not found to be induced by Wnt/β-catenin signaling in HMECs, neither in the context of TGF-β signaling nor its absence. It may be that *GSC* induction in HMECs is complicated by a need for such intermediaries. The orthologs of Siamois and Twin in mouse and human have not yet been identified. These genes are essential for Spemann organizer formation and *Gsc* induction in *Xenopus* (Fan and Sokol 1997; Nishita et al. 2000). I did find putative binding sites in the human *GSC* promoter for proteins that partner with β-catenin (tcf/lefs)to activate transcription, however, by genome sequence analysis. This makes the observed lack of *GSC* induction in HMECs, even by a constitutively active, functional Tcf/Lef protein, a lingering mystery. My analysis of signaling upstream of *GSC* is summarized in Figure 2A.

Signaling downstream of Goosecoid

I have found that Gsc expression generates changes in the levels of the adhesion components E-cadherin, N-cadherin, various catenins, and fibronectin, cytoskeleton-related proteins vimentin, cytokeratins, and IQGAP1, transcription factors *SIP1, Snail*, and FOXC2, and the PDGF β-type ligand and receptor duo (Figure 2B). I do not yet know which of these genes are direct transcriptional targets of Gsc and which are further downstream. Gsc could potentially directly regulate any of these, or could elicit the observed cellular phenotypes by acting through other gene(s) altogether. The exact mechanism by which Gsc induces these phenotypes and gene expression changes thus remains unclear.

I employed a bioinformatics approach in an attempt to identify genes that Gsc is especially likely to directly control. The promoter sequences of certain genes, including *IQGAP1* and *N-cadherin*, were subsequently found to contain putative Gsc-binding sequences. In fact, this is how we came to discover that Gsc upregulates the levels of IQGAP1 in human mammary cells (HMECs). This analysis will eventually be complemented by an unbiased screen for direct transcriptional targets of Gsc in HMECs using chIP-on-chip technology in collaboration with others in our lab and the Young lab at MIT (Wyrick and Young 2002). This study will hopeful reveal useful insights into the signaling downstream to Gsc and the other EMT-inducing transcription factors.

Gsc might induce an EMT by merely regulating just one or two genes, as this might be enough to destabilize the epithelial phenotype and initiate a cascade of EMT changes. Another mechanism by which Gsc might trigger an EMT is by inducing a TGF-β autocrine loop or other autocrine signaling. The induction of a TGF-β autocrine loop is known to stabilize the EMT state in certain systems, such that the mesenchymal state can be maintained independently of exogenous factors (Grunert et al. 2003). I do not yet know whether such a loop is induced in HMECs expressing ectopic Gsc, although my very preliminary data suggests that such a loop is probably not essential for the induction of the EMT at least.

I also do not know whether the genes whose levels are altered by Gsc expression are specific to this transcription factor or whether they are merely characteristic of an induced EMT state. For example, components of the PDGF pathway are known to be upregulated when epithelial cells undergo a TGF-β induced EMT (Jechlinger et al. 2003; Jechlinger et al. 2006). Very recently, PDGFR signaling has also been shown to have a role in metastasis and in the maintenance of the EMT(Jechlinger et al. 2006). These discoveries indicate that the upregulation

of PDGFR signaling components in HMECs is most likely not a Gsc-specific effect, unless in these published systems Gsc mediates the induction and maintenance of an EMT by TGF-β.

Goosecoid implicates other Spemann organizer genes in cancer metastasis

This thesis work, together with the work of others, suggests that the process of metastasis is not quite as complex as once thought. A tumor cell may well be able to simultaneously acquire several cellular capabilities, such as the loss of cell anchorage, resistance to anoikis, motility and invasiveness, by reactivating a developmental gene program for which it is already internally hardwired.

Gsc was originally known for its essential role in normal *Xenopus* cell migration in the organizer at gastrulation (Steinbeisser et al. 1995; Ferreiro et al. 1998; Yao and Kessler 2001). I therefore hypothesized that genes driving the organizer phenotype are likely to be especially potent in promoting cell migration in cancer metastasis, and my data supports this hypothesis. Other genes that mediate organizer biology might likewise become active during cancer progression. In addition to Gsc, the known organizer-specific transcription factors include Xlim1, Orthodenticle 2 (Otx2), Xnot, Siamois, Twin and several members of the forkhead family including Foxa2 and Foxd3 (Lemaire and Kodjabachian 1996; Koide et al. 2005). Other genes that are specific to the organizer but do not encode transcription factors include noggin, chordin, and follistatin (De Robertis 1995; Lemaire and Kodjabachian 1996).

Perspectives

Taken together, the present results implicate the Spemann organizer gene, *Goosecoid*, in tumor metastasis. Moreover, they suggest that the re-activation of conserved organizer genes to

promote tumor invasion and metastasis may be a recurrent theme in human cancer. My findings therefore warrant a comprehensive examination of these genes in multiple types of human malignancies.

Figure 1. Proposed mechanisms of *Goosecoid* induction

A (far left). Gsc is induced in a non-cell autonomous fashion as a result of contextual signals received by cells just at the invasion front of a tumor. The induction of Gsc in the context of the pre-existed background of accumulated genetic and epigenetic lesions induces a transdifferentiation (or EMT) in these cells. These cells are thus now especially aggressive. Alternatively, Gsc could be similarly induced by contextual signals but might not generate an EMT until additional, cooperative genetic lesions were sustained. This mechanism of Gsc induction is non-cell autonomous, and the sustained expression of Gsc may be dependent on continued exposure to stromal signals.

A (far right). Gsc is induced in abnormal tissue as a results of a genetic or epigenetic stable event. Gsc may immediately induce a phenotypic effect or it may demonstrate a functional cellular effect until additional collaborating mutations are incurred.

B. Gsc is expressed in normal cells present at low numbers within a given tissue. Upon the transformation and proliferation of such a cell, the levels of Gsc will effectively appear upregulated because the original Gsc-expressing cells are now more prevalent within the tissue. In this paradigm, Gsc is acting cell-autonomously and may or may not provide a functional effect.

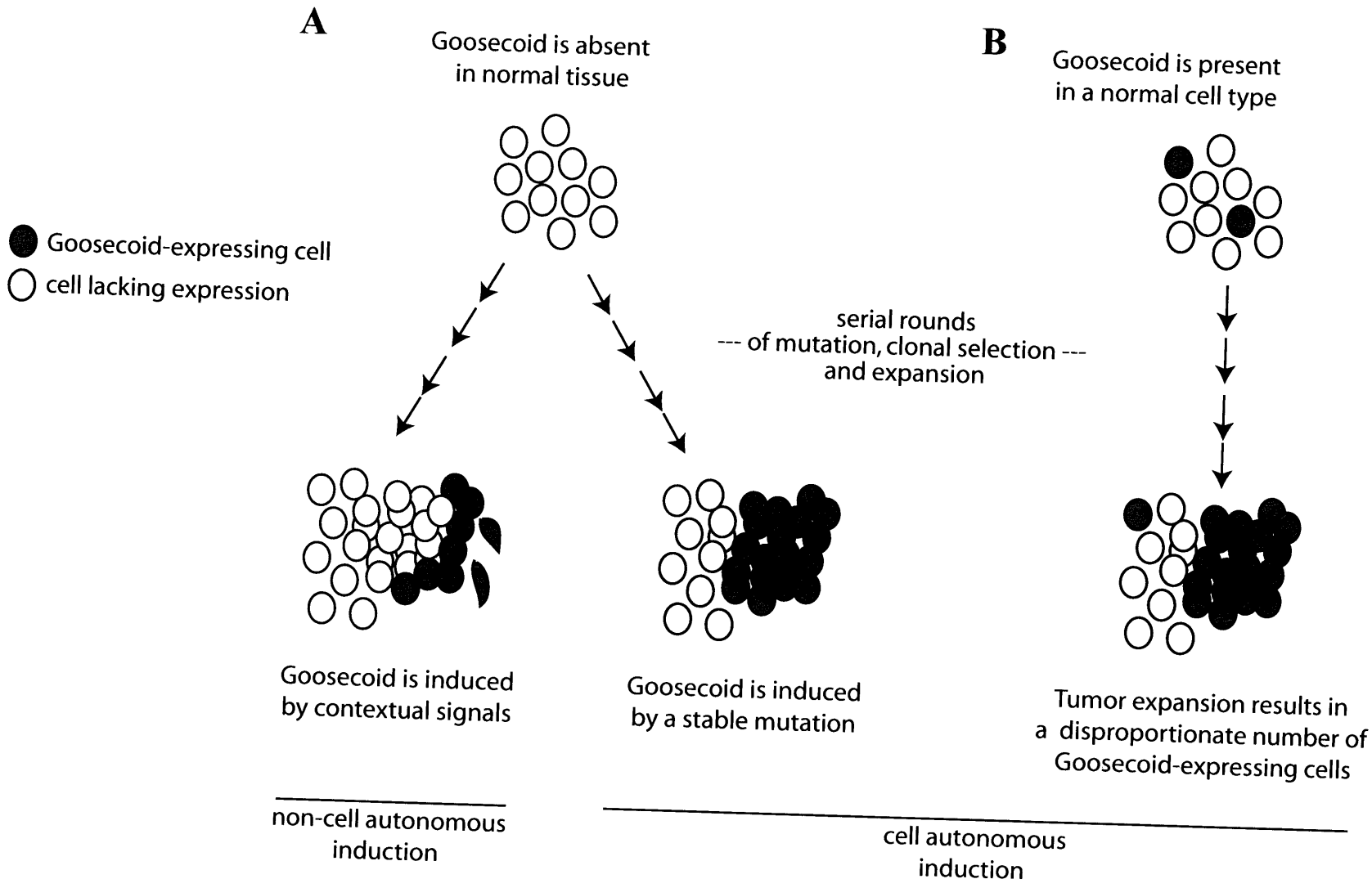
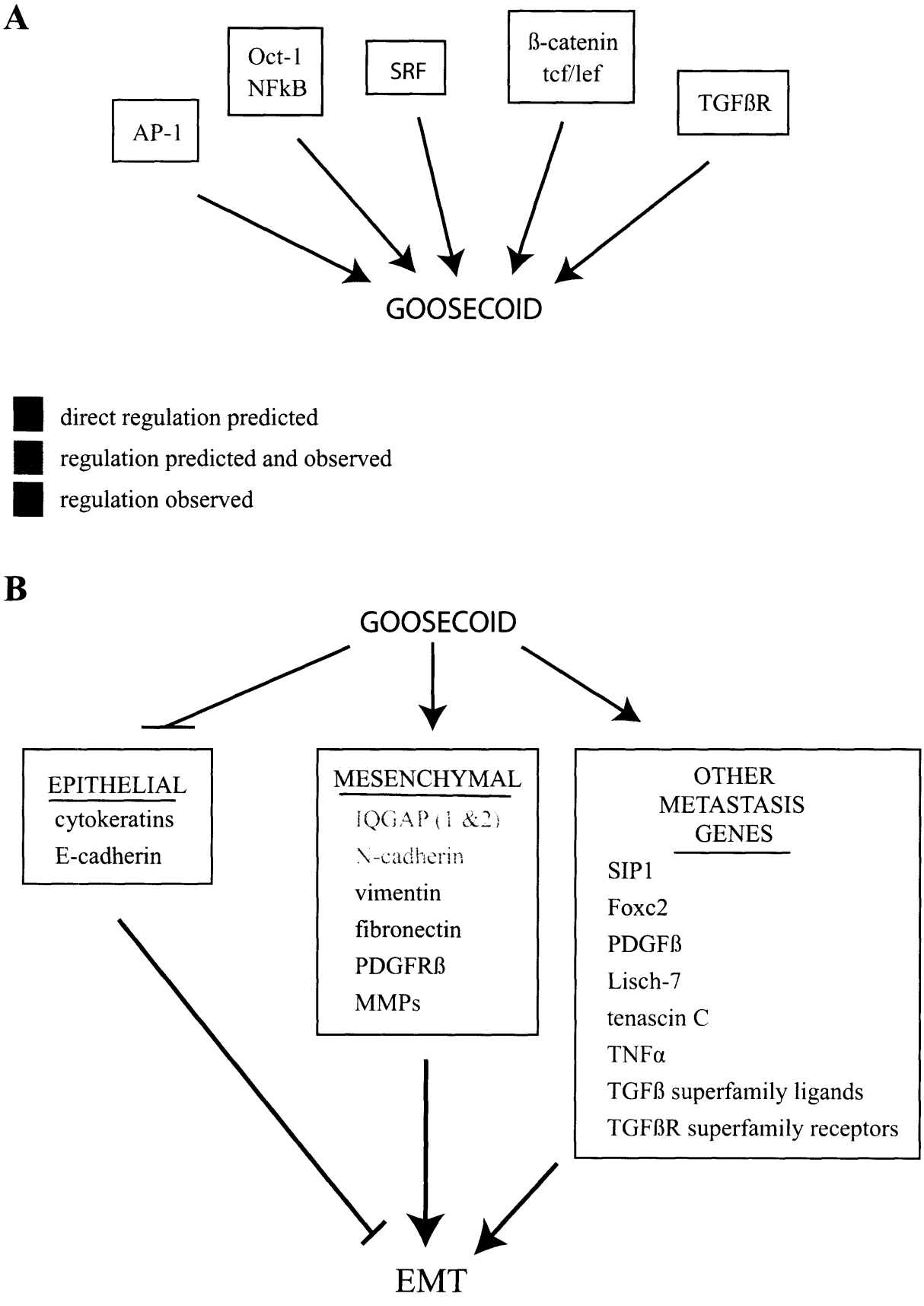

A
Goosecoid is absent
in normal tissue
Goosecoid-expressing cell
cell lacking expression
serial rounds
--- of mutation, clonal selection ---
and expansion
Goosecoid is induced
by contextual signals
Goosecoid is induced
by a stable mutation
non-cell autonomous
induction
B
Goosecoid is present
in a normal cell type
Tumor expansion results in
a disproportionate number of
Goosecoid-expressing cells
cell autonomous
induction

Figure 2. Cell signaling upstream and downstream of *Goosecoid*

A. A summary of the putative and confirmed signals found to be upstream of *Gsc* is shown in the schematic. The signals are color-coded according to whether they were confirmed to indeed upregulate *Gsc* expression or whether they are predicted by my bioinformatics analysis.

B. A summary of the putative and confirmed signals found to be downstream of Gsc is shown in the schematic. The signals are color-coded according to whether they were confirmed to indeed upregulate Gsc expression or whether they are predicted by my bioinformatics analysis. Gsc promotes an EMT, and thus tumor invasion and metastasis, by presumably antagonizing the expression of certain epithelial markers while upregulating certain mesenchymal markers. Additional genes already implicated in the EMT and/or metastasis are also shown.


A
AP-1
Oct-1
NFkB
SRF
ß-catenin
tcf/lef
TGFßR
GOOSECOID
direct regulation predicted
regulation predicted and observed
regulation observed
B
GOOSECOID
EPITHELIAL
cytokeratins
E-cadherin
MESENCHYMAL
IQGAP (1 &2)
N-cadherin
vimentin
fibronectin
PDGFRß
MMPs
OTHER
METASTASIS
GENES
SIP1
Foxc2
PDGFß
Lisch-7
tenascin C
TNFα
TGFß superfamily ligands
TGFßR superfamily receptors
EMT

Appendix

Invasive Carcinoma Cells Respond to Contextual Signals *in vivo*

Kimberly A. Hartwell[1,2], Tan Ince[1,3], and Robert A. Weinberg[1,2]

[1]Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142 USA
[2]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA
[3]Department of Pathology, Harvard Medical School, Brigham and Women's Hospital, Boston, MA 02115 USA

to date, published in

Weinberg R A (2007) *The Biology of Cancer* (Garland Science, New York).

The creation and characterization of the BPLER tumor model was the work of Dr. Tan Ince, as were the pilot vimentin stainings of tumor cells *in situ*. Kimberly Hartwell stained adjacent tumor sections for additional EMT markers and carefully examined the resulting stains for an inverse pattern of mesenchymal and epithelial marker staining. These observations inspired the double fluorescent stainings which were also done by Kimberly Hartwell.

Abstract

An epithelial-mesenchymal transition (EMT) is known to promote the invasive and metastatic capabilities of cancer cells. Moreover, an EMT may represent one mechanism by which contextual signals from surrounding stromal cells influence the phenotype of specific subpopulations of carcinoma cells within a tumor mass. A demonstration of an EMT transdifferentiation event occurring in an *in vivo* setting as a result of exposure to contextual signals has not been observed. Here, I have examined human breast cancer xenograft tumors for evidence of an EMT transdifferentiation event taking place *in vivo*. BPLER tumor cells, which are vimentin negative and cytokeratin positive at injection, generate tumors containing tumor cells that are vimentin positive and cytokeratin negative. This change is specific to the invasion front of these tumors and constitutes evidence of an EMT transition occurring *in vivo*, presumably as a result of exposure to contextual signals emanating from the surrounding stroma.

Introduction

Tumor cells are thought to acquire metastatic capabilities by way of the gene expression dictated by the combined effects of tumor cell of origin, genetic background, the assorted genetic and epigenetic lesions incurred during the process of tumor progression, as well as contextual signals (Fearon and Vogelstein 1990; Bissell and Radisky 2001; Hynes 2003; Gupta et al. 2005a). Carcinoma cells at the leading edge, or invasion front, of tumors are in close juxtaposition with their stromal cell neighbors. Whereas tumor cells at the edge of a tumor are thus exposed to stromal signals and can undergo heterotypic interactions with stromal cells, cells at the interior of a tumor are shielded from the effects of stromal contact. The role of the stroma surrounding tumors in promoting tumor progression is increasingly well described (Bissell and Radisky 2001). Moreover, certain signals observed to be elevated at the tumor invasion front are known correlates of metastasis, such as Tenascin C and nuclear β-catenin, suggesting that cells in this subregion of a tumor may be especially aggressive compared to cells at the tumor center (Jahkola et al. 1998; Brabletz et al. 2001).

Evidence of E-cadherin downregulation has been observed specifically at the invasive edge of certain clinical colon tumors (Brabletz et al. 2001). E-cadherin is a hallmark of the epithelial-mesenchymal transition (EMT), a transdifferentiation process that allows epithelial cells to acquire mesenchymal properties such as a spindly morphology and motility and invasiveness during normal processes such as development and wound healing. This transition can be induced in tumor cells *in vitro* and indeed correlates with the invasive and metastatic phenotype (Thiery 2002). However, much controversy exists as to whether this transdifferentiation event actual occurs *in vivo* within a growing tumor.

The reduced E-cadherin levels at the edge of colon tumors suggested two provocative possibilities- that an EMT might indeed occur in tumors and that contextual signals might be capable of inducing this transdifferentiation event (Brabletz et al. 2001). What had yet to be observed, however, to the best of our knowledge, was evidence of a coordinated switching between the epithelial and mesenchymal states in a given subpopulation of tumor cells. Such an observation would strongly argue that a single tumor cell subpopulation could indeed undergo a switch from an epithelial to mesenchymal phenotype and that such a transdifferentiation likely results from exposure to adjacent stroma.

As I was examining a potential role for Gsc in tumor metastasis, my colleague, Dr. Tan Ince, a postdoc in the Weinberg lab, was examining the effects of cell-of-origin differences on the biology and prognosis of breast cancers. He developed a human tumor xenograft model, denoted BPLER, that faithfully recapitulates human ductal adenocarcinomas and their associated clinical biology (unpublished results, T. A. Ince et al., manuscript in preparation). Provocatively, unlike tumors generated using human cells isolated in a traditional way that did not allow for the modeling of clinically-relevant tumors, the tumors in his model demonstrated metastatic capabilities *in vivo*. One of the many significant implications of this work was that both acquired genetic lesions and cell-of-origin differences dictated the biology, including the metastatic propensity, of human tumors.

Tan made the side observation that cells within these xenograft tumors showed evidence of an edge effect. He observed by immunoperoxidase immunohistochemical staining that vimentin, a mesenchymal cytoskeletal protein, was upregulated at the tumor invasion front. Given our overlapping interests, I examined these tumors more closely for evidence of a coordinated EMT shift.

Results

By immunoperoxidase staining, I observed the same upregulation of vimentin at the invasion front of the BPLER tumors (Figure 1A). I also stained adjacent tumor sections for other EMT-associated markers in parallel, including E-cadherin, β-catenin, and cytokeratins. Strikingly, the cytokeratin staining pattern suggested a complementary downregulation of cytokeratins in tumor regions that showed an upregulation in adjacent tumor sections (Figure 1B). Cells at the invasive edge showed almost a complete absence of cytokeratin stain whereas cells at the tumor center stained strongly. In contrast, no difference in the staining pattern of E-cadherin or β-catenin across tumor subregions was visible. Instead, both proteins were consistently detected at the membrane, where they are known to comprise adherens junctions (data not shown).

To confirm that a single population of tumor cells was showing evidence of simultaneous vimentin upregulation and cytokeratin downregulation, I completed a immunohistochemical double staining of the same tumors using fluorescently-tagged secondary antibodies. These stainings confirmed that the tumor cell subpopulation in contact with adjacent stroma showed evidence of downregulated cytokeratins and upregulated vimentin (Figure 1C).

Tan confirmed by western blotting that the originally injected, polyclonal BPLER cell populations was vimentin negative (data not shown), indicating that vimentin expression was acquired by a subset of these cells *in vivo*.

We also noted that there was no evidence of a morphological difference between cells at the tumor edge versus the center. Also, the described transdifferentiation was not evidenced at all invasive edges of BPLER tumors, and we do not yet understand the source of this variation within a given tumor.

Discussion

Together, these observations demonstrate that tumor cells can indeed undergo changes in gene expression as a result of contact with stromal cell neighbors. We were probably only able to observe the vimentin edge effect because the antibody that we used was specific for human, not mouse, vimentin. Consequently, we could be sure that the vimentin positive cells detected were injected human tumor cells rather than leaving open the possibility that they were merely adjacent mouse fibroblasts, which are known to express vimentin.

Moreover, the change that the cells at the invasive edge were undergoing was specifically an EMT transdifferentiation, a phenomenon that remains quite controversial among clinical cancer biologists despite the numerous *in vitro* and *in vivo* experimental studies linking the EMT and tumor metastasis (Thiery and Sleeman 2006). We do not yet understand why these given EMT markers were changing yet other markers such as E-cadherin were not. It may be that the partial EMTs observed experimentally have as much, if not more, clinical relevance than do the complete EMT shifts that are sometimes observed instead.

Finally, the BPLER xenograft model provides the means for further study of the EMT phenomenon *in vivo*, given that it is a tractable experimental system.

Figure 1. Invasive carcinoma cells respond to contextual signals *in vivo*

A. The tumor cells that comprise the edge of human mammary tumor xenografts in mice express vimentin (brown), whereas the cells at the center of the tumors do not. Immunoperoxidase staining was used to detect vimentin by way of a human-specific vimentin antibody.

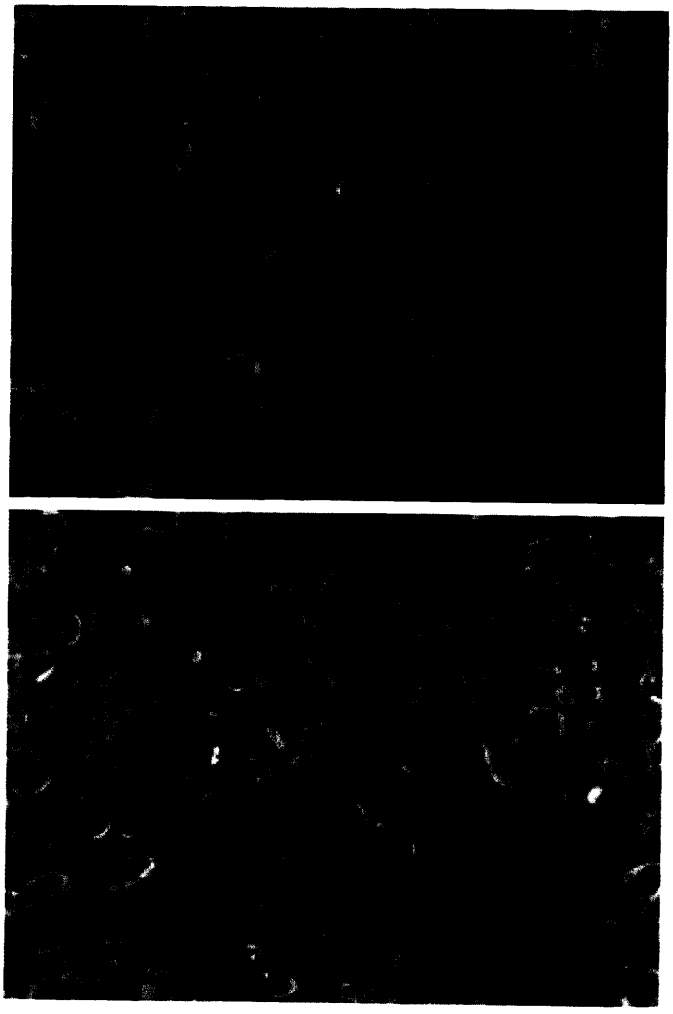
B. Conversely, cells localized to the edge of the same xenograft tumors examined in panel A display a downregulation of cytokeratin expression (brown). Expression of this epithelial cell marker is maintained in the cells at the tumor interior, as evidenced by immunoperoxidase staining.

C. Double staining of a single tumor section for both vimentin and cytokeratins confirms the hypothesized edge effect. Experimentally transformed human mammary epithelial cells form a tumor in which the bulk of the carcinoma cells express epithelial cytokeratins (red). Cancer cells that are in contact with the surrounding mouse stroma (blue) at the edge of the tumor have shut down keratin expression and instead have induced vimentin expression (green). This shift, or transdifferentiation, at the tumor edge is likely occurring in response to contextual stromal signals.

A

B

C



Materials and Methods

Immunohistochemical detection of vimentin, cytokeratins, E-cadherin and β-catenin was performed either using the avidin : biotinylated enzyme complex procedure (Vector Laboratories, Inc.) or using fluorophore-labeled secondary antibodies (Molecular Probes). Formalin-fixed, paraffin-embedded tissue sections were probed using human-specific mouse monoclonal vimentin antibody (#NCL-vim-v9, Novacastra), pankeratin-specific rabbit antibody (#PU071-UP, Biogenex), and the E-cadherin and β-catenin antibodies detailed in the main methods section of this thesis. Antigen retrieval was performed using heat treatment in sodium citrate buffer.

References

Al-Hajj, M., M.S. Wicha, A. Benito-Hernandez, S.J. Morrison, and M.F. Clarke. 2003. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A 100: 3983-8.

Aoki, M., A. Hecht, U. Kruse, R. Kemler, and P.K. Vogt. 1999. Nuclear endpoint of Wnt signaling: neoplastic transformation induced by transactivating lymphoid-enhancing factor 1. Proc Natl Acad Sci U S A 96: 139-44.

Arendt, D., U. Technau, and J. Wittbrodt. 2001. Evolution of the bilaterian larval foregut. Nature 409: 81-5.

Aslakson, C.J. and F.R. Miller. 1992. Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor. Cancer Res 52: 1399-405.

Ataliotis, P., K. Symes, M.M. Chou, L. Ho, and M. Mercola. 1995. PDGF signalling is required for gastrulation of Xenopus laevis. Development 121: 3099-110.

Baldus, S.E., S.P. Monig, S. Huxel, S. Landsberg, F.G. Hanisch, K. Engelmann, P.M. Schneider, J. Thiele, A.H. Holscher, and H.P. Dienes. 2004. MUC1 and nuclear beta-catenin are coexpressed at the invasion front of colorectal carcinomas and are both correlated with tumor prognosis. Clin Cancer Res 10: 2790-6.

Barth, A.I., A.L. Pollack, Y. Altschuler, K.E. Mostov, and W.J. Nelson. 1997. NH2-terminal deletion of beta-catenin results in stable colocalization of mutant beta-catenin with adenomatous polyposis coli protein and altered MDCK cell adhesion. J Cell Biol 136: 693-706.

Batlle, E., E. Sancho, C. Franci, D. Dominguez, M. Monfar, J. Baulida, and A. Garcia De Herreros. 2000. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. Nat Cell Biol 2: 84-9.

Beiter, K., E. Hiendlmeyer, T. Brabletz, F. Hlubek, A. Haynl, C. Knoll, T. Kirchner, and A. Jung. 2005. beta-Catenin regulates the expression of tenascin-C in human colorectal tumors. Oncogene 24: 8200-4.

Bernards, R. and R.A. Weinberg. 2002. A progression puzzle. Nature 418: 823.

Berx, G. and F. Van Roy. 2001. The E-cadherin/catenin complex: an important gatekeeper in breast cancer tumorigenesis and malignant progression. Breast Cancer Res 3: 289-93.

Birchmeier, C., W. Birchmeier, and B. Brand-Saberi. 1996. Epithelial-mesenchymal transitions in cancer progression. Acta Anat 156: 217-26.

Birchmeier, W. and C. Birchmeier. 1995. Epithelial-mesenchymal transitions in development and tumor progression. Exs 74: 1-15.

Bissell, M.J. and D. Radisky. 2001. Putting tumours in context. Nat Rev Cancer 1: 46-54.

Blum, M., E.M. De Robertis, T. Kojis, C. Heinzmann, I. Klisak, D. Geissert, and R.S. Sparkes. 1994. Molecular cloning of the human homeobox gene goosecoid (GSC) and mapping of the gene to human chromosome 14q32.1. Genomics 21: 388-93.

Blum, M., S.J. Gaunt, K.W. Cho, H. Steinbeisser, B. Blumberg, D. Bittner, and E.M. De Robertis. 1992. Gastrulation in the mouse: the role of the homeobox gene goosecoid. Cell 69: 1097-106.

Blumberg, B., C.V. Wright, E.M. De Robertis, and K.W. Cho. 1991. Organizer-specific homeobox genes in Xenopus laevis embryos. Science 253: 194-6.

Bolos, V., H. Peinado, M.A. Perez-Moreno, M.F. Fraga, M. Esteller, and A. Cano. 2003. The transcription factor Slug represses E-cadherin expression and induces epithelial to

mesenchymal transitions: a comparison with Snail and E47 repressors. J Cell Sci 116: 499-511.
Brabletz, T., A. Jung, S. Dag, F. Hlubek, and T. Kirchner. 1999. beta-catenin regulates the expression of the matrix metalloproteinase-7 in human colorectal cancer. Am J Pathol 155: 1033-8.
Brabletz, T., A. Jung, K. Hermann, K. Gunther, W. Hohenberger, and T. Kirchner. 1998. Nuclear overexpression of the oncoprotein beta-catenin in colorectal cancer is localized predominantly at the invasion front. Pathol Res Pract 194: 701-4.
Brabletz, T., A. Jung, S. Reu, M. Porzner, F. Hlubek, L.A. Kunz-Schughart, R. Knuechel, and T. Kirchner. 2001. Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment. Proc Natl Acad Sci U S A 98: 10356-61.
Brabletz, T., A. Jung, S. Spaderna, F. Hlubek, and T. Kirchner. 2005. Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression. Nat Rev Cancer 5: 744-9.
Brennan, J., C.C. Lu, D.P. Norris, T.A. Rodriguez, R.S. Beddington, and E.J. Robertson. 2001. Nodal signalling in the epiblast patterns the early mouse embryo. Nature 411: 965-9.
Briggs, M.W. and D.B. Sacks. 2003. IQGAP proteins are integral components of cytoskeletal regulation. EMBO Rep 4: 571-4.
Broun, M., S. Sokol, and H.R. Bode. 1999. Cngsc, a homologue of goosecoid, participates in the patterning of the head, and is expressed in the organizer region of Hydra. Development 126: 5245-54.
Cano, A., M.A. Perez-Moreno, I. Rodrigo, A. Locascio, M.J. Blanco, M.G. del Barrio, F. Portillo, and M.A. Nieto. 2000. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. Nat Cell Biol 2: 76-83.
Carpenter, A.E., T.R. Jones, M.R. Lamprecht, C. Clarke, I.H. Kang, O. Friman, D.A. Guertin, J.H. Chang, R.A. Lindquist, J. Moffat, P. Golland, and D.M. Sabatini. 2006. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol 7: R100.
Chambers, A.F., A.C. Groom, and I.C. MacDonald. 2002. Dissemination and growth of cancer cells in metastatic sites. Nat Rev Cancer 2: 563-72.
Chen, G. and A.J. Courey. 2000. Groucho/TLE family proteins and transcriptional repression. Gene 249: 1-16.
Cho, K.W., B. Blumberg, H. Steinbeisser, and E.M. De Robertis. 1991. Molecular nature of Spemann's organizer: the role of the Xenopus homeobox gene goosecoid. Cell 67: 1111-20.
Christiansen, J.J. and A.K. Rajasekaran. 2006. Reassessing epithelial to mesenchymal transition as a prerequisite for carcinoma invasion and metastasis. Cancer Res 66: 8319-26.
Clark, E.A., T.R. Golub, E.S. Lander, and R.O. Hynes. 2000. Genomic analysis of metastasis reveals an essential role for RhoC. Nature 406: 532-5.
Comijn, J., G. Berx, P. Vermassen, K. Verschueren, L. van Grunsven, E. Bruyneel, M. Mareel, D. Huylebroeck, and F. van Roy. 2001. The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. Mol Cell 7: 1267-78.
Dalal, B.I., P.A. Keown, and A.H. Greenberg. 1993. Immunocytochemical localization of secreted transforming growth factor-beta 1 to the advancing edges of primary tumors and to lymph node metastases of human mammary carcinoma. Am J Pathol 143: 381-9.

Danilov, V., M. Blum, A. Schweickert, M. Campione, and H. Steinbeisser. 1998. Negative autoregulation of the organizer-specific homeobox gene goosecoid. J Biol Chem 273: 627-35.

de Iongh, R.U., E. Wederell, F.J. Lovicu, and J.W. McAvoy. 2005. Transforming growth factor-beta-induced epithelial-mesenchymal transition in the lens: a model for cataract formation. Cells Tissues Organs 179: 43-55.

De Robertis, E.M. 1995. Developmental biology. Dismantling the organizer. Nature 374: 407-8.

Debnath, J. and J.S. Brugge. 2005. Modelling glandular epithelial cancers in three-dimensional cultures. Nat Rev Cancer 5: 675-88.

DeRobertis, E.M. 2004. Goosecoid and Gastrulation. In Gastrulation, From Cells to Embryo (ed. C.D. Stern). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Dvorak, H.F. 1986. Tumors: wounds that do not heal. Similarities between tumor stroma generation and wound healing. N Engl J Med 315: 1650-9.

Eccles, S.A., C. Box, and W. Court. 2005. Cell migration/invasion assays and their application in cancer drug discovery. Biotechnol Annu Rev 11: 391-421.

Eger, A., K. Aigner, S. Sonderegger, B. Dampier, S. Oehler, M. Schreiber, G. Berx, A. Cano, H. Beug, and R. Foisner. 2005. DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. Oncogene 24: 2375-85.

Elbashir, S.M., J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, and T. Tuschl. 2001. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature 411: 494-8.

Elenbaas, B., L. Spirio, F. Koerner, M.D. Fleming, D.B. Zimonjic, J.L. Donaher, N.C. Popescu, W.C. Hahn, and R.A. Weinberg. 2001. Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. Genes Dev 15: 50-65.

Elenbaas, B. and R.A. Weinberg. 2001. Heterotypic signaling between epithelial tumor cells and fibroblasts in carcinoma formation. Exp Cell Res 264: 169-84.

Fan, M.J. and S.Y. Sokol. 1997. A role for Siamois in Spemann organizer formation. Development 124: 2581-9.

Fata, J.E., Z. Werb, and M.J. Bissell. 2004. Regulation of mammary gland branching morphogenesis by the extracellular matrix and its remodeling enzymes. Breast Cancer Res 6: 1-11.

Fearon, E.R. and B. Vogelstein. 1990. A genetic model for colorectal tumorigenesis. Cell 61: 759-67.

Ferreiro, B., M. Artinger, K. Cho, and C. Niehrs. 1998. Antimorphic goosecoids. Development 125: 1347-59.

Fidler, I.J. 2002. The organ microenvironment and cancer metastasis. Differentiation 70: 498-505.

Fidler, I.J. 2003. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. Nat Rev Cancer 3: 453-8.

Fidler, I.J. and M.L. Kripke. 1977. Metastasis results from preexisting variant cells within a malignant tumor. Science 197: 893-5.

Filosa, S., J.A. Rivera-Perez, A.P. Gomez, A. Gansmuller, H. Sasaki, R.R. Behringer, and S.L. Ang. 1997. Goosecoid and HNF-3beta genetically interact to regulate neural tube patterning during mouse embryogenesis. Development 124: 2843-54.

Folkman, J. 1986. How is blood vessel growth regulated in normal and neoplastic tissue? G.H.A. Clowes memorial Award lecture. Cancer Res 46: 467-73.

Frew, S.E.A. 2004. Investigation of the role of IQGAP1 in metastasis. In Institute Thesis Archives. Massachusetts Institute of Technology, Cambridge.
Frisch, S.M. and H. Francis. 1994. Disruption of epithelial cell-matrix interactions induces apoptosis. J Cell Biol 124: 619-26.
Funke, B., B. Saint-Jore, A. Puech, H. Sirotkin, L. Edelmann, C. Carlson, S. Raft, R.K. Pandita, R. Kucherlapati, A. Skoultchi, and B.E. Morrow. 1997. Characterization and mutation analysis of goosecoid-like (GSCL), a homeodomain-containing gene that maps to the critical region for VCFS/DGS on 22q11. Genomics 46: 364-72.
Galili, N., H.S. Baldwin, J. Lund, R. Reeves, W. Gong, Z. Wang, B.A. Roe, B.S. Emanuel, S. Nayak, C. Mickanin, M.I. Budarf, and C.A. Buck. 1997. A region of mouse chromosome 16 is syntenic to the DiGeorge, velocardiofacial syndrome minimal critical region. Genome Res 7: 399.
Galili, N., J.A. Epstein, I. Leconte, S. Nayak, and C.A. Buck. 1998. Gscl, a gene within the minimal DiGeorge critical region, is expressed in primordial germ cells and the developing pons. Dev Dyn 212: 86-93.
Germain, S., M. Howell, G.M. Esslemont, and C.S. Hill. 2000. Homeodomain and winged-helix transcription factors recruit activated Smads to distinct promoter elements via a common Smad interaction motif. Genes Dev 14: 435-51.
Gilbert, S.F. 1997. Developmental Biology. In, pp. 209-252. Sinauer Associates, Inc., Sunderland, MA.
Gilles, C., M. Polette, P. Birembaut, N. Brunner, and E.W. Thompson. 1997. Expression of c-ets-1 mRNA is associated with an invasive, EMT-derived phenotype in breast carcinoma cell lines. Clin Exp Metastasis 15: 519-26.
Gilles, C., M. Polette, M. Mestdagt, B. Nawrocki-Raby, P. Ruggeri, P. Birembaut, and J.M. Foidart. 2003. Transactivation of vimentin by beta-catenin in human breast cancer cells. Cancer Res 63: 2658-64.
Going, J.J., T.J. Anderson, S. Battersby, and C.C. MacIntyre. 1988. Proliferative and secretory activity in human breast during natural and artificial menstrual cycles. Am J Pathol 130: 193-204.
Gradl, D., M. Kuhl, and D. Wedlich. 1999. The Wnt/Wg signal transducer beta-catenin controls fibronectin expression. Mol Cell Biol 19: 5576-87.
Grunert, S., M. Jechlinger, and H. Beug. 2003. Diverse cellular and molecular mechanisms contribute to epithelial plasticity and metastasis. Nat Rev Mol Cell Biol 4: 657-65.
Gumbiner, B.M. 1996. Cell adhesion: the molecular basis of tissue architecture and morphogenesis. Cell 84: 345-57.
Gupta, P.B., C. Kuperwasser, J.P. Brunet, S. Ramaswamy, W.L. Kuo, J.W. Gray, S.P. Naber, and R.A. Weinberg. 2005a. The melanocyte differentiation program predisposes to metastasis after neoplastic transformation. Nat Genet 37: 1047-54.
Gupta, P.B., S. Mani, J. Yang, K. Hartwell, and R.A. Weinberg. 2005b. The evolving portrait of cancer metastasis. Cold Spring Harb Symp Quant Biol 70: 291-7.
Hajra, K.M., D.Y. Chen, and E.R. Fearon. 2002. The SLUG zinc-finger protein represses E-cadherin in breast cancer. Cancer Res 62: 1613-8.
Hanahan, D. and R.A. Weinberg. 2000. The hallmarks of cancer. Cell 100: 57-70.
Harland, R. and J. Gerhart. 1997. Formation and function of Spemann's organizer. Annu Rev Cell Dev Biol 13: 611-67.

Haub, O. and M. Goldfarb. 1991. Expression of the fibroblast growth factor-5 gene in the mouse embryo. Development 112: 397-406.
Hay, E.D. and A. Zuk. 1995. Transformations between epithelium and mesenchyme: normal, pathological, and experimentally induced. Am J Kidney Dis 26: 678-90.
Heasman, J., A. Crawford, K. Goldstone, P. Garner-Hamrick, B. Gumbiner, P. McCrea, C. Kintner, C.Y. Noro, and C. Wylie. 1994. Overexpression of cadherins and underexpression of beta-catenin inhibit dorsal mesoderm induction in early Xenopus embryos. Cell 79: 791-803.
Hebert, J.M., M. Boyle, and G.R. Martin. 1991. mRNA localization studies suggest that murine FGF-5 plays a role in gastrulation. Development 112: 407-15.
Hemmati-Brivanlou, A. and D.A. Melton. 1992. A truncated activin receptor inhibits mesoderm induction and formation of axial structures in Xenopus embryos. Nature 359: 609-14.
Holbro, T., G. Civenni, and N.E. Hynes. 2003. The ErbB receptors and their role in cancer progression. Exp Cell Res 284: 99-110.
Holmgren, L., M.S. O'Reilly, and J. Folkman. 1995. Dormancy of micrometastases: balanced proliferation and apoptosis in the presence of angiogenesis suppression. Nat Med 1: 149-53.
Hoschuetzky, H., H. Aberle, and R. Kemler. 1994. Beta-catenin mediates the interaction of the cadherin-catenin complex with epidermal growth factor receptor. J Cell Biol 127: 1375-80.
Hu, M., J. Yao, L. Cai, K.E. Bachman, F. van den Brule, V. Velculescu, and K. Polyak. 2005. Distinct epigenetic changes in the stromal cells of breast cancers. Nat Genet 37: 899-905.
Huber, M.A., N. Azoitei, B. Baumann, S. Grunert, A. Sommer, H. Pehamberger, N. Kraut, H. Beug, and T. Wirth. 2004. NF-kappaB is essential for epithelial-mesenchymal transition and metastasis in a model of breast cancer progression. J Clin Invest 114: 569-81.
Huber, M.A., N. Kraut, and H. Beug. 2005. Molecular requirements for epithelial-mesenchymal transition during tumor progression. Curr Opin Cell Biol 17: 548-58.
Hynes, R.O. 2003. Metastatic potential: generic predisposition of the primary tumor or rare, metastatic variants-or both? Cell 113: 821-3.
Jahkola, T., T. Toivonen, I. Virtanen, K. von Smitten, S. Nordling, K. von Boguslawski, C. Haglund, H. Nevanlinna, and C. Blomqvist. 1998. Tenascin-C expression in invasion border of early breast cancer: a predictor of local and distant recurrence. Br J Cancer 78: 1507-13.
Jechlinger, M., S. Grunert, I.H. Tamir, E. Janda, S. Ludemann, T. Waerner, P. Seither, A. Weith, H. Beug, and N. Kraut. 2003. Expression profiling of epithelial plasticity in tumor progression. Oncogene 22: 7155-69.
Jechlinger, M., A. Sommer, R. Moriggl, P. Seither, N. Kraut, P. Capodiecci, M. Donovan, C. Cordon-Cardo, H. Beug, and S. Grunert. 2006. Autocrine PDGFR signaling promotes mammary cancer metastasis. J Clin Invest 116: 1561-70.
Jimenez, G., C.P. Verrijzer, and D. Ish-Horowicz. 1999. A conserved motif in goosecoid mediates groucho-dependent repression in Drosophila embryos. Mol Cell Biol 19: 2080-7.
Jouanneau, J., J. Gavrilovic, D. Caruelle, M. Jaye, G. Moens, J.P. Caruelle, and J.P. Thiery. 1991. Secreted or nonsecreted forms of acidic fibroblast growth factor produced by transfected epithelial cells influence cell morphology, motility, and invasive potential. Proc Natl Acad Sci U S A 88: 2893-7.
Kang, Y., P.M. Siegel, W. Shu, M. Drobnjak, S.M. Kakonen, C. Cordon-Cardo, T.A. Guise, and J. Massague. 2003. A multigenic program mediating breast cancer metastasis to bone. Cancer Cell 3: 537-49.

Kikuchi, A. 2000. Regulation of beta-catenin signaling in the Wnt pathway. Biochem Biophys Res Commun 268: 243-8.

Kitadai, Y., L.M. Ellis, S.L. Tucker, G.F. Greene, C.D. Bucana, K.R. Cleary, Y. Takahashi, E. Tahara, and I.J. Fidler. 1996. Multiparametric in situ mRNA hybridization analysis to predict disease recurrence in patients with colon carcinoma. Am J Pathol 149: 1541-51.

Koide, T., T. Hayata, and K.W. Cho. 2005. Xenopus as a model system to study transcriptional regulatory networks. Proc Natl Acad Sci U S A 102: 4943-8.

Korinek, V., N. Barker, P.J. Morin, D. van Wichen, R. de Weger, K.W. Kinzler, B. Vogelstein, and H. Clevers. 1997. Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. Science 275: 1784-7.

Ku, M., S.Y. Sokol, J. Wu, M.I. Tussie-Luna, A.L. Roy, and A. Hata. 2005. Positive and negative regulation of the transforming growth factor beta/activin target gene goosecoid by the TFII-I family of transcription factors. Mol Cell Biol 25: 7144-57.

Kuniyasu, H., P. Troncoso, D. Johnston, C.D. Bucana, E. Tahara, I.J. Fidler, and C.A. Pettaway. 2000. Relative expression of type IV collagenase, E-cadherin, and vascular endothelial growth factor/vascular permeability factor in prostatectomy specimens distinguishes organ-confined from pathologically advanced prostate cancers. Clin Cancer Res 6: 2295-308.

Labbe, E., C. Silvestri, P.A. Hoodless, J.L. Wrana, and L. Attisano. 1998. Smad2 and Smad3 positively and negatively regulate TGF beta-dependent transcription through the forkhead DNA-binding protein FAST2. Mol Cell 2: 109-20.

Ladanyi, A., F. Sipos, D. Szoke, O. Galamb, B. Molnar, and Z. Tulassay. 2006. Laser microdissection in translational and clinical research. Cytometry A 69: 947-60.

Larabell, C.A., M. Torres, B.A. Rowning, C. Yost, J.R. Miller, M. Wu, D. Kimelman, and R.T. Moon. 1997. Establishment of the dorso-ventral axis in Xenopus embryos is presaged by early asymmetries in beta-catenin that are modulated by the Wnt signaling pathway. J Cell Biol 136: 1123-36.

Laurent, M.N., I.L. Blitz, C. Hashimoto, U. Rothbacher, and K.W. Cho. 1997. The Xenopus homeobox gene twin mediates Wnt induction of goosecoid in establishment of Spemann's organizer. Development 124: 4905-16.

Lemaire, P. and L. Kodjabachian. 1996. The vertebrate organizer: structure and molecules. Trends Genet 12: 525-31.

Leyns, L., T. Bouwmeester, S.H. Kim, S. Piccolo, and E.M. De Robertis. 1997. Frzb-1 is a secreted antagonist of Wnt signaling expressed in the Spemann organizer. Cell 88: 747-56.

Li, L., J. Roden, B.E. Shapiro, B.J. Wold, S. Bhatia, S.J. Forman, and R. Bhatia. 2005. Reproducibility, fidelity, and discriminant validity of mRNA amplification for microarray analysis from primary hematopoietic cells. J Mol Diagn 7: 48-56.

Lin, E.Y., J.G. Jones, P. Li, L. Zhu, K.D. Whitney, W.J. Muller, and J.W. Pollard. 2003. Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. Am J Pathol 163: 2113-26.

Lin, S.Y., W. Xia, J.C. Wang, K.Y. Kwong, B. Spohn, Y. Wen, R.G. Pestell, and M.C. Hung. 2000. Beta-catenin, a novel prognostic marker for breast cancer: its roles in cyclin D1 expression and cancer progression. Proc Natl Acad Sci U S A 97: 4262-6.

Liu, P., M. Wakamiya, M.J. Shea, U. Albrecht, R.R. Behringer, and A. Bradley. 1999. Requirement for Wnt3 in vertebrate axis formation. Nat Genet 22: 361-5.

Ma, X.J., R. Salunga, J.T. Tuggle, J. Gaudet, E. Enright, P. McQuary, T. Payette, M. Pistone, K. Stecker, B.M. Zhang, Y.X. Zhou, H. Varnholt, B. Smith, M. Gadd, E. Chatfield, J. Kessler,

T.M. Baer, M.G. Erlander, and D.C. Sgroi. 2003. Gene expression profiles of human breast cancer progression. Proc Natl Acad Sci U S A 100: 5974-9.

Ma, X.J., Z. Wang, P.D. Ryan, S.J. Isakoff, A. Barmettler, A. Fuller, B. Muir, G. Mohapatra, R. Salunga, J.T. Tuggle, Y. Tran, D. Tran, A. Tassin, P. Amon, W. Wang, E. Enright, K. Stecker, E. Estepa-Sabal, B. Smith, J. Younger, U. Balis, J. Michaelson, A. Bhan, K. Habin, T.M. Baer, J. Brugge, D.A. Haber, M.G. Erlander, and D.C. Sgroi. 2004. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. Cancer Cell 5: 607-16.

Menard, S., S. Fortis, F. Castiglioni, R. Agresti, and A. Balsari. 2001. HER2 as a prognostic factor in breast cancer. Oncology 61 Suppl 2: 67-72.

Minn, A.J., G.P. Gupta, P.M. Siegel, P.D. Bos, W. Shu, D.D. Giri, A. Viale, A.B. Olshen, W.L. Gerald, and J. Massague. 2005. Genes that mediate breast cancer metastasis to lung. Nature 436: 518-24.

Moll, R., M. Mitze, U.H. Frixen, and W. Birchmeier. 1993. Differential loss of E-cadherin expression in infiltrating ductal and lobular breast carcinomas. Am J Pathol 143: 1731-42.

Moody, S.E., D. Perez, T.C. Pan, C.J. Sarkisian, C.P. Portocarrero, C.J. Sterner, K.L. Notorfrancesco, R.D. Cardiff, and L.A. Chodosh. 2005. The transcriptional repressor Snail promotes mammary tumor recurrence. Cancer Cell 8: 197-209.

Moon, R.T. and D. Kimelman. 1998. From cortical rotation to organizer gene expression: toward a molecular explanation of axis specification in Xenopus. Bioessays 20: 536-45.

Nicolson, G.L. 1988. Cancer metastasis: tumor cell and host organ properties important in metastasis to specific secondary sites. Biochim Biophys Acta 948: 175-224.

Niehrs, C. 2004. Regionally specific induction by the Spemann-Mangold organizer. Nat Rev Genet 5: 425-34.

Niehrs, C., R. Keller, K.W. Cho, and E.M. De Robertis. 1993. The homeobox gene goosecoid controls cell migration in Xenopus embryos. Cell 72: 491-503.

Nishita, M., M.K. Hashimoto, S. Ogata, M.N. Laurent, N. Ueno, H. Shibuya, and K.W. Cho. 2000. Interaction between Wnt and TGF-beta signalling pathways during formation of Spemann's organizer. Nature 403: 781-5.

Oft, M., K.H. Heider, and H. Beug. 1998. TGFbeta signaling is necessary for carcinoma cell invasiveness and metastasis. Curr Biol 8: 1243-52.

Okada, H., T.M. Danoff, R. Kalluri, and E.G. Neilson. 1997. Early role of Fsp1 in epithelial-mesenchymal transformation. Am J Physiol 273: F563-74.

Ondruschka, C., P. Buhtz, C. Motsch, B. Freigang, R. Schneider-Stock, A. Roessner, and C. Boltze. 2002. Prognostic value of MMP-2, -9 and TIMP-1,-2 immunoreactive protein at the invasive front in advanced head and neck squamous cell carcinomas. Pathol Res Pract 198: 509-15.

Orimo, A., P.B. Gupta, D.C. Sgroi, F. Arenzana-Seisdedos, T. Delaunay, R. Naeem, V.J. Carey, A.L. Richardson, and R.A. Weinberg. 2005. Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. Cell 121: 335-48.

Perez-Moreno, M.A., A. Locascio, I. Rodrigo, G. Dhondt, F. Portillo, M.A. Nieto, and A. Cano. 2001. A new role for E12/E47 in the repression of E-cadherin expression and epithelial-mesenchymal transitions. J Biol Chem 276: 27424-31.

Perl, A.K., P. Wilgenbus, U. Dahl, H. Semb, and G. Christofori. 1998. A causal role for E-cadherin in the transition from adenoma to carcinoma. Nature 392: 190-3.

Portella, G., S.A. Cumming, J. Liddell, W. Cui, H. Ireland, R.J. Akhurst, and A. Balmain. 1998. Transforming growth factor beta is essential for spindle cell conversion of mouse skin carcinoma in vivo: implications for tumor invasion. Cell Growth Differ 9: 393-404.

Porter, D., J. Lahti-Domenici, A. Keshaviah, Y.K. Bae, P. Argani, J. Marks, A. Richardson, A. Cooper, R. Strausberg, G.J. Riggins, S. Schnitt, E. Gabrielson, R. Gelman, and K. Polyak. 2003. Molecular markers in ductal carcinoma in situ of the breast. Mol Cancer Res 1: 362-75.

Postigo, A.A. and D.C. Dean. 2000. Differential expression and function of members of the zfh-1 family of zinc finger/homeodomain repressors. Proc Natl Acad Sci U S A 97: 6391-6.

Price, J.E., A. Polyzos, R.D. Zhang, and L.M. Daniels. 1990. Tumorigenicity and metastasis of human breast carcinoma cell lines in nude mice. Cancer Res 50: 717-21.

Prindull, G. 2005. Hypothesis: cell plasticity, linking embryonal stem cells to adult stem cell reservoirs and metastatic cancer cells? Exp Hematol 33: 738-46.

Ramaswamy, S., K.N. Ross, E.S. Lander, and T.R. Golub. 2003. A molecular signature of metastasis in primary solid tumors. Nat Genet 33: 49-54.

Rasbridge, S.A., C.E. Gillett, S.A. Sampson, F.S. Walsh, and R.R. Millis. 1993. Epithelial (E-) and placental (P-) cadherin cell adhesion molecule expression in breast carcinoma. J Pathol 169: 245-50.

Reya, T., S.J. Morrison, M.F. Clarke, and I.L. Weissman. 2001. Stem cells, cancer, and cancer stem cells. Nature 414: 105-11.

Rivera-Perez, J.A., M. Mallo, M. Gendron-Maguire, T. Gridley, and R.R. Behringer. 1995. Goosecoid is not an essential component of the mouse gastrula organizer but is required for craniofacial and rib development. Development 121: 3005-12.

Ronnov-Jessen, L., O.W. Petersen, and M.J. Bissell. 1996. Cellular changes involved in conversion of normal to malignant breast: importance of the stromal reaction. Physiol Rev 76: 69-125.

Rosivatz, E., I. Becker, K. Specht, E. Fricke, B. Luber, R. Busch, H. Hofler, and K.F. Becker. 2002. Differential expression of the epithelial-mesenchymal transition regulators snail, SIP1, and twist in gastric cancer. Am J Pathol 161: 1881-91.

Sakai, D., Y. Tanaka, Y. Endo, N. Osumi, H. Okamoto, and Y. Wakamatsu. 2005. Regulation of Slug transcription in embryonic ectoderm by beta-catenin-Lef/Tcf and BMP-Smad signaling. Dev Growth Differ 47: 471-82.

Sasai, Y., B. Lu, H. Steinbeisser, D. Geissert, L.K. Gont, and E.M. De Robertis. 1994. Xenopus chordin: a novel dorsalizing factor activated by organizer-specific homeobox genes. Cell 79: 779-90.

Sasaki, H. and B.L. Hogan. 1993. Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo. Development 118: 47-59.

Savagner, P., K.M. Yamada, and J.P. Thiery. 1997. The zinc-finger protein slug causes desmosome dissociation, an initial and necessary step for growth factor-induced epithelial-mesenchymal transition. J Cell Biol 137: 1403-19.

Schmidt-Kittler, O., T. Ragg, A. Daskalakis, M. Granzow, A. Ahr, T.J. Blankenstein, M. Kaufmann, J. Diebold, H. Arnholdt, P. Muller, J. Bischoff, D. Harich, G. Schlimok, G. Riethmuller, R. Eils, and C.A. Klein. 2003. From latent disseminated cells to overt metastasis: genetic analysis of systemic breast cancer progression. Proc Natl Acad Sci U S A 100: 7737-42.

Shook, D. and R. Keller. 2003. Mechanisms, mechanics and function of epithelial-mesenchymal transitions in early development. Mech Dev 120: 1351-83.
Slager, H.G., K.A. Lawson, A.J. van den Eijnden-van Raaij, S.W. de Laat, and C.L. Mummery. 1991. Differential localization of TGF-beta 2 in mouse preimplantation and early postimplantation development. Dev Biol 145: 205-18.
Smith, H.S., S.R. Wolman, and A.J. Hackett. 1984. The biology of breast cancer at the cellular level. Biochim Biophys Acta 738: 103-23.
Sporn, M.B. 1996. The war on cancer. Lancet 347: 1377-81.
Steeg, P.S. 2006. Tumor metastasis: mechanistic insights and clinical challenges. Nat Med 12: 895-904.
Steinbeisser, H., A. Fainsod, C. Niehrs, Y. Sasai, and E.M. De Robertis. 1995. The role of gsc and BMP-4 in dorsal-ventral patterning of the marginal zone in Xenopus: a loss-of-function study using antisense RNA. Embo J 14: 5230-43.
Sternlicht, M.D., A. Lochter, C.J. Sympson, B. Huey, J.P. Rougier, J.W. Gray, D. Pinkel, M.J. Bissell, and Z. Werb. 1999. The stromal proteinase MMP3/stromelysin-1 promotes mammary carcinogenesis. Cell 98: 137-46.
Stewart, S.A., D.M. Dykxhoorn, D. Palliser, H. Mizuno, E.Y. Yu, D.S. An, D.M. Sabatini, I.S. Chen, W.C. Hahn, P.A. Sharp, R.A. Weinberg, and C.D. Novina. 2003. Lentivirus-delivered stable gene silencing by RNAi in primary cells. Rna 9: 493-501.
Sun, X., E.N. Meyers, M. Lewandoski, and G.R. Martin. 1999. Targeted disruption of Fgf8 causes failure of cell migration in the gastrulating mouse embryo. Genes Dev 13: 1834-46.
Tetsu, O. and F. McCormick. 1999. Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. Nature 398: 422-6.
Thiery, J.P. 2002. Epithelial-mesenchymal transitions in tumour progression. Nat Rev Cancer 2: 442-54.
Thiery, J.P. 2003. Epithelial-mesenchymal transitions in development and pathologies. Curr Opin Cell Biol 15: 740-6.
Thiery, J.P. and J.P. Sleeman. 2006. Complex networks orchestrate epithelial-mesenchymal transitions. Nat Rev Mol Cell Biol 7: 131-42.
Timmerman, L.A., J. Grego-Bessa, A. Raya, E. Bertran, J.M. Perez-Pomares, J. Diez, S. Aranda, S. Palomo, F. McCormick, J.C. Izpisua-Belmonte, and J.L. de la Pompa. 2004. Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. Genes Dev 18: 99-115.
Van Dyke, T. and T. Jacks. 2002. Cancer modeling in the modern era: progress and challenges. Cell 108: 135-44.
Vega, S., A.V. Morales, O.H. Ocana, F. Valdes, I. Fabregat, and M.A. Nieto. 2004. Snail blocks the cell cycle and confers resistance to cell death. Genes Dev 18: 1131-43.
Voleti, B. and A. Agrawal. 2005. Regulation of basal and induced expression of C-reactive protein through an overlapping element for OCT-1 and NF-kappaB on the proximal promoter. J Immunol 175: 3386-90.
Wang, L.H. 2004. Molecular signaling regulating anchorage-independent growth of cancer cells. Mt Sinai J Med 71: 361-7.
Wang, W., S. Goswami, K. Lapidus, A.L. Wells, J.B. Wyckoff, E. Sahai, R.H. Singer, J.E. Segall, and J.S. Condeelis. 2004. Identification and testing of a gene expression signature of invasive carcinoma cells within primary mammary tumors. Cancer Res 64: 8585-94.

Watabe, T., S. Kim, A. Candia, U. Rothbacher, C. Hashimoto, K. Inoue, and K.W. Cho. 1995. Molecular mechanisms of Spemann's organizer formation: conserved growth factor synergy between Xenopus and mouse. Genes Dev 9: 3038-50.

Weeraratna, A.T. 2005. Discovering causes and cures for cancer from gene expression analysis. Ageing Res Rev 4: 548-63.

Weigand, R.A., W.M. Isenberg, J. Russo, M.J. Brennan, and M.A. Rich. 1982. Blood vessel invasion and axillary lymph node involvement as prognostic indicators for human breast cancer. Cancer 50: 962-9.

Weigelt, B., J.L. Peterse, and L.J. van 't Veer. 2005. Breast cancer metastasis: markers and models. Nat Rev Cancer 5: 591-602.

Weiss, L. 1990. Metastatic inefficiency. Adv Cancer Res 54: 159-211.

Wieser, R., J.L. Wrana, and J. Massague. 1995. GS domain mutations that constitutively activate T beta R-I, the downstream signaling component in the TGF-beta receptor complex. Embo J 14: 2199-208.

Wilson, D., G. Sheng, T. Lecuit, N. Dostatni, and C. Desplan. 1993. Cooperative dimerization of paired class homeo domains on DNA. Genes Dev 7: 2120-34.

Wyckoff, J.B., J.G. Jones, J.S. Condeelis, and J.E. Segall. 2000. A critical step in metastasis: in vivo analysis of intravasation at the primary tumor. Cancer Res 60: 2504-11.

Wyrick, J.J. and R.A. Young. 2002. Deciphering gene expression regulatory networks. Curr Opin Genet Dev 12: 130-6.

Xu, L., S. Begum, J.D. Hearn, and R.O. Hynes. 2006. GPR56, an atypical G protein-coupled receptor, binds tissue transglutaminase, TG2, and inhibits melanoma tumor growth and metastasis. Proc Natl Acad Sci U S A 103: 9023-8.

Xue, C., D. Plieth, C. Venkov, C. Xu, and E.G. Neilson. 2003. The gatekeeper effect of epithelial-mesenchymal transition regulates the frequency of breast cancer metastasis. Cancer Res 63: 3386-94.

Yamada, G., A. Mansouri, M. Torres, E.T. Stuart, M. Blum, M. Schultz, E.M. De Robertis, and P. Gruss. 1995. Targeted mutation of the murine goosecoid gene results in craniofacial defects and neonatal death. Development 121: 2917-22.

Yang, J., S.A. Mani, J.L. Donaher, S. Ramaswamy, R.A. Itzykson, C. Come, P. Savagner, I. Gitelman, A. Richardson, and R.A. Weinberg. 2004. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell 117: 927-39.

Yao, J. and D.S. Kessler. 2001. Goosecoid promotes head organizer activity by direct repression of Xwnt8 in Spemann's organizer. Development 128: 2975-87.

Zavadil, J., L. Cermak, N. Soto-Nieves, and E.P. Bottinger. 2004. Integration of TGF-beta/Smad and Jagged1/Notch signalling in epithelial-to-mesenchymal transition. Embo J 23: 1155-65.

Zhou, B.P., J. Deng, W. Xia, J. Xu, Y.M. Li, M. Gunduz, and M.C. Hung. 2004. Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition. Nat Cell Biol 6: 931-40.